Représentations de groupes de Lie et fonctionnement géométrique du cerveau

Alexandre Afgoustidis

Université Paris VII - Denis Diderot

École Doctorale de Sciences Mathématiques de Paris Centre

Thèse de doctorat

Discipline : Mathématiques

présentée par

Alexandre Afgoustidis

Représentations de groupes de Lie et fonctionnement géométrique du cerveau

---

dirigée par Daniel Bennequin

Soutenue le 4 juillet 2016 devant le jury composé de :

M. Daniel Bennequin	Université Paris VII	directeur
Mme Giovanna Citti	Université de Bologne	rapporteur
M. Michel Duflo	Université Paris VII	examinateur
M. Jean Petitot	EHESS	examinateur
M. François Rouvière	Université de Nice	rapporteur
M. Alessandro Sarti	CNRS et EHESS	examinateur
M. Georges Skandalis	Université Paris VII	examinateur
M. David A. Vogan, Jr.	MIT	rapporteur

UMR7586
Institut de Mathématiques de Jussieu-
Paris Rive Gauche
Université Paris Diderot
Campus des Grands Moulins
Bâtiment Sophie Germain, case 7012
75205 Paris Cedex 13

Ecole Doctorale de Sciences
Mathématiques de Paris Centre
4 place Jussieu
75252 Paris Cedex 05
Case 188

Résumé

Cette thèse étudie deux problèmes indépendants où l'analyse harmonique non-commutative, théorie des représentations de groupes de Lie, joue un rôle.

Le premier problème a été suggéré par les neurosciences à Daniel Bennequin, qui me l'a proposé. Nous partons des profils récepteurs des neurones du cortex visuel primaire, et de la géométrie de la répartition des spécialités des neurones à la surface de cette aire corticale. Nous rappelons comment de remarquables propriétés des cartes d'orientation qu'on trouve dans cette aire corticale peuvent être reproduites par les tirages de champs aléatoires gaussiens invariants en loi dont les tirages explorent un facteur de Plancherel de la représentation quasi-régulière du groupe euclidien sur l'espace des fonctions sur le plan euclidien. Nous signalons que cette interprétation permet de construire, sur la sphère et sur le plan hyperbolique, des structures géométriques qui rappellent (qualitativement et quantitativement) les cartes d'orientation du cortex visuel. L'intervention naturelle d'autres groupes dans le traitement des informations sensorielles et la préparation du mouvement nous invite à envisager la construction d'objets analogues pour d'autres espaces homogènes comme une question mathématique d'intérêt indépendant. Nous étudions alors les champs aléatoires gaussiens invariants sur les espaces homogènes riemanniens, en donnant des constructions explicites issues de la théorie des représentations (d'après Akiva N. Yaglom) et en démontrant que dans une unité de volume adaptée, la mesure géométrique moyenne de l'ensemble des zéros ne dépend que des dimensions de la source et du but.

Le second problème a été suggéré par George W. Mackey en 1975 ; il est interne à la théorie des représentations de groupes de Lie réductifs réels. Nous décrivons une bijection entre le dual tempéré d'un groupe de Lie linéaire connexe réductif et le dual unitaire de son groupe de déplacements de Cartan. Le second groupe est une contraction du premier, au sens d'Inonü et Wigner ; afin de comprendre la bijection précédente à l'aide de la notion de contraction, nous utilisons, pour toute représentation tempérée irréductible du premier groupe, une famille d'opérateurs de contraction qui permet de suivre le comportement des vecteurs lisses (dans une réalisation géométrique adaptée) au cours de la contraction et d'observer leur convergence vers la représentation du second groupe indiquée par notre bijection. Nous utilisons ensuite nos résultats pour donner une nouvelle preuve de la conjecture de Connes-Kasparov pour les groupes de Lie linéaires connexes réductifs, en suivant la méthode utilisée par Nigel Higson en 2008.

Le manuscrit contient par ailleurs le récit d'une étude sur des données expérimentales issues d'enregistrements électrophysiologiques de l'activité de cellules de Purkinje dans le cervelet vestibulaire de rats vigiles, à l'aide d'éléments matriciels de représentations unitaires du groupe de Galilée.

Mots-clefs

• Représentations de groupes de Lie, analyse harmonique non-commutative ;
• Cortex visuel primaire, cartes d'orientation, densité des pinwheels, champs aléatoires gaussiens sur les espaces homogènes ;
• Système vestibulaire et vestibulo-cervelet, groupe de Galilée, éléments matriciels de représentations unitaires ;
• Groupes de Lie réductifs, dual tempéré, contractions de groupes de Lie, groupe de déplacements de Cartan, isomorphisme de Connes-Kasparov.

Invariant Harmonic Analysis and Geometry in the Workings of the Brain

Abstract

Two independent problems will be considered in this thesis ; both feature non-commutative harmonic analysis—in other words, Lie group representation theory.

The first problem was suggested by recent results from Neuroscience to Daniel Bennequin, who suggested it to me in turn. Our starting points are the receptive profiles of neurons in the primary visual cortex, and the geometrical properties of the maps describing how the neurons' specializations are distributed on the surface of that part of the cortex. We recall that some remarkable properties of the "orientation maps" to be found in that cortical area are strikingly well reproduced by typical samples from Gaussian random fields on the Euclidean plane whose probability distribution is invariant under the Euclidean group, and whose samples probe an irreducible Plancherel factor of the quasi-regular representation of the Euclidean group on the space of functions on the Euclidean plane. We indicate that this interpretation makes it possible to build geometrical structures which call to mind, both qualitatively and qualitatively, the orientation maps of the visual cortex. Because the intervention of other groups is natural in the study of sensory information and motion planning, the construction of analogous structures on more general homogeneous spaces has an independent interest ; we thus proceed to a study of invariant Gaussian random fields on riemannian homogeneous spaces, outline explicit constructions based on group representation theory (following Akiva N. Yaglom), and prove that when expressed an appropriate volume unit, the geometric measure of the zero-set of an invariant field depends only on the dimensions of the source and target spaces.

The second problem was suggested by an old question from George W. Mackey and recent work by Nigel Higson ; the question was asked in 1975, and is internal to the representation theory of real reductive Lie groups. We describe a bijection between the tempered dual of a linear connected reductive Lie group and the unitary dual of its Cartan motion group. The second group is a contraction of the first, in the terminology of Inonü and Wigner ; in order to understand our bijection in terms of contractions of Lie groups, we consider an arbitrary irreducible tempered representation of the first group and introduce a family of contraction operators which make it possible to follow the individual smooth vectors (in an appropriate geometric realization for the representation) during the contraction ; we observe their convergence to the members of a carrier space for the representation of the Cartan motion group indicated by our bijection. We then use our results to give a new proof of the Connes-Kasparov conjecture in the case of linear connected reductive Lie groups, extending relatively recent work by Nigel Higson.

This manuscript also contains a report on an attempt to use some matrix elements of unitary representations of the Galilei group in the study of electrophysiological recordings of the activity of Purkinje cells in the vestibulocerebellum of live (and alert) rats.

Key words

- Representation theory of Lie groups, non-commutative harmonic analysis ;
- Primary visual cortex, orientation maps, pinwheel density, Gaussian random fields on homogeneous spaces ;
- Vestibular system, cerebellum, Galilei group, matrix elements of unitary representations ;
- Reductive Lie groups, tempered dual, Lie group contractions, Cartan motion group, Connes-Kasparov isomorphism.

Table des matières

Chapitre 0

Introduction (en français)

Contents

Les trois parties de cette thèse sont largement indépendantes. La première s'appuie sur l'organisation du cortex visuel primaire des mammifères et contient des résultats d'analyse harmonique invariante et de probabilités ; la deuxième contient le récit d'une exploration (hélas infructueuse) de données expérimentales sur le fonctionnement du cervelet des rats. La troisième s'occupe d'un problème sur les représentations tempérées de groupes de Lie réductifs réels, et de ses conséquences pour la K-théorie des $C^\star$ algèbres – ce qui est bien loin du cerveau. Cette introduction, en plus de présenter les résultats de ma thèse, doit donc expliquer ce que ces problèmes ont de commun ; le texte de chacune des parties ne dira plus ces liens. L'introduction sera longue.

Vous êtes en train de parcourir une thèse de mathématiques. Elle porte sur les symétries, sur la théorie des groupes, sur celle de leurs représentations. J'y essaie notamment, c'est de là que vient le titre, de signaler quelques situations où la théorie des représentations peut nous aider à parler de ce que fait le cerveau ; mais l'essentiel de son contenu est bien sûr dans les résultats mathématiques que je vais démontrer. Je ne voudrais pas que les remarques de biologie qui ouvriront mon texte fassent oublier que la troisième partie de ma thèse, où j'étudie les analogies entre le dual tempéré d'un groupe de Lie réel réductif et le dual unitaire de son groupe de déplacements de Cartan, y est tout sauf marginale : c'est là que sont les résultats qui m'ont demandé le plus de travail.

L'introduction va d'abord décrire quelques-unes des structures biologiques qui nous permettent d'entendre, de voir, et de percevoir et planifier nos mouvements. Si l'on veut bien y chercher des groupes, on les verra, discrètement pour l'instant – il n'est pas étonnant de les rencontrer, le groupe des translations de la droite réelle pour l'audition, celui des translations et ceux des déplacements du plan et de l'espace pour l'organisation de la population des neurones visuels, le groupe de Galilée pour la perception du mouvement. Une partie de ma thèse tournera autour du fait que l'analyse harmonique invariante permet de mettre au jour le rôle des symétries dans les modèles qui permettent de parler de ces structures et du fait que le contenu mathématique de ces modèles peut être adapté à des symétries différentes au moyen de la théorie des groupes et de celle des représentations.

Je voudrais qu'il soit clair que convoquer la théorie des représentations de groupes de Lie pour nous aider à parler de problèmes pratiques est une idée ancienne ; que cette idée est couverte de grands succès, et qu'elle a joué un rôle essentiel dans la naissance et dans l'essor de la théorie elle-même. Pour cela, je rappellerai dans la deuxième partie de l'introduction plusieurs épisodes de l'histoire du sujet. Ce sera l'occasion de présenter certains des thèmes de la théorie des représentations qui sont au coeur de mon travail, et d'introduire quelques-unes des notions qui joueront les premiers rôles par la suite.

Alors, et alors seulement, il sera temps de parler des résultats de ma thèse. On me pardonnera, j'espère, la longueur de la discussion générale : le chapitre d'exploration de données expérimentales aurait pu à lui seul justifier la pertinence du point de vue qui donne son titre à mon manuscrit. Mais il n'y suffira pas ; c'est cette introduction qui doit vous en convaincre.

1 Perception, profils récepteurs, cartes corticales

Lorsque nous nous déplaçons ou lorsque des modifications surviennent autour de nous, l'information qui parvient aux structures biologiques chargées de nous renseigner sur notre environnement change aussi. Pas de répit, peu de constance pour nos capteurs sensoriels.

Mais montez dans un bus, et traversez une place aux pavés vieillissants : le sol et les statues ne vous paraîtront pas trembler. Repassant le soir au même endroit, les visages de pierre ne vous sembleront pas avoir changé depuis le matin ; le mouvement qui les rapproche ou les éloigne de vous, et le fait qu'ils soient maintenant encadrés d'obscurité plutôt que de soleil, ne vous empêcheront pas de les reconnaître.

La stabilité de l'image que nous nous faisons de ce qui nous entoure dépend de notre capacité à repérer les modifications dans notre environnement, à distinguer celles qui dépendent de nos actions de celles qui n'en dépendent pas. Sans moyen efficace de reconstituer notre déplacement relativement aux sources, sans moyen de gérer le contexte dans lequel s'insèrent les objets autour de nous (par exemple de s'adapter à une luminosité qui change), d'anticiper correctement et d'utiliser ensemble invariance et anticipation pour agir, nous serions perdus : pas d'identification cohérente des objets, de notre place dans le monde, peut-être pas de notion stable d'espace [1].

Le cerveau *doit* savoir mettre de l'ordre géométrique dans les informations qu'il reçoit, inventer des structures d'invariance correspondantes, s'en servir pour prévoir et pour agir. Et bien sûr, il le fait, sans cesse.

Il le fait si bien que nous oublions que bouger les yeux change complètement ce que captent les bâtonnets et cônes, que fixer un objet lorsqu'on court ou qu'on joue nécessite un contrôle délicat des mouvements relatifs de la tête et des yeux. Ce n'est qu'à la faveur d'un instant de fatigue ou de rêverie que nous sentons le monde reculer quand notre train avance, que nous voyons coupée en deux une branche dont nous sépare le bois de la fenêtre.

Nous sommes à l'aise pour gérer ces difficultés ou ces ambiguïtés, qui sont toutes de nature géométrique et dont aucune n'est triviale. Nous sommes tellement à l'aise que nous ne les gérons pas consciemment, nous n'avons pas (sauf exception) à y *penser*. Beaucoup de résultats récents des neurosciences semblent indiquer que *l'architecture même* de plusieurs aires cérébrales est plus qu'adaptée pour gérer ces ambiguïtés, et qu'il n'est pas absurde de dire que ces aires traitent l'information *de façon plus ou moins directement invariante* pour des structures abstraites qui gouvernent les ambiguïtés géométriques ; celle de groupe est très commode lorsqu'elle est adaptée à la situation, bien que ce soit loin d'être toujours le cas.

Je vais présenter quelques-unes des structures biologiques qui permettent au cerveau d'appréhender certains stimuli sensoriels. Les parties de ma thèse qui concernent le cerveau essaient de dire comment la théorie des groupes et celle de leurs représentations permettent d'essayer de commencer à parler de ces structures, des modèles qui cherchent à les comprendre, et souvent, de généraliser ces modèles.

Ma thèse s'occupe de mathématiques et les faits expérimentaux que je vais présenter ici servent simplement de motivation à certaines de ses parties ; je donnerai peu de détails anatomiques et biologiques. Il faut cependant en donner quelques-uns pour comprendre d'où viennent les modèles mathématiques qui seront mon point de départ, et pourquoi ils correspondent à des problèmes importants pour le cerveau. Ces détails rapides sont donnés en pensant à un(e) mathématicien(ne) qui voudrait lire ma thèse et comprendre la motivation des résultats qu'on y trouve ; je prie celles et ceux de mes lecteurs qui connaissent le détail des systèmes biologiques concernés de pardonner les énormes raccourcis. Les notes

1. Poincaré le disait mieux il y a un siècle, j'y reviendrai.

de bas de page signalent des remords, mais elles n'apporteront que peu de précisions.

1.1 Comment notre oreille perçoit les sons

Commençons par observer un schéma de l'oreille et des structures situées derrière notre tympan.

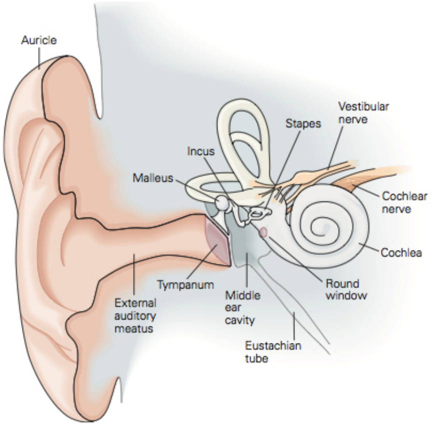


Figure 1 [31].

La partie "haute" de l'oreille interne (en vert léger ci-contre) est l'organe dont nous nous servons pour enregistrer les mouvements de notre tête ; ce sera l'un des acteurs principaux du chapitre 6, et j'en parlerai dans la section 1.3. de cette introduction.

Mais regardons d'abord la partie de l'oreille interne qui transmet les sons au cerveau. Ce que je vais dire du système auditif reprend une partie de la présentation de A. J. Hupsteth dans le traité coordonné par E. Kandel et J. Schwartz ([31], chapitre 30). Les figures en sont extraites.

C'est à l'intérieur de la *cochlée*, l'organe tubulaire en forme d'escargot de la figure 1, que le premier nerf auditif prend sa source. La cochlée est longue de trois bons centimètres, et comporte trois chambres remplies de liquide, comme sur la figure 2 ; la chambre haute [2] et la chambre basse communiquent au bout [3] de la cochlée, et le liquide (l'*endolymphe*) peut circuler librement de l'une à l'autre. La chambre médiane est close, et sa base est la *membrane basilaire* ; c'est elle qui abrite les capteurs connectés aux neurones qui transmettent l'information sonore au cerveau [4].

Les trois osselets [5] convertissent les vibrations de l'air et du tympan en mouvements du liquide des chambres haute et basse : le dernier osselet frappe à l'entrée de la chambre haute, et met en mouvement l'endolymphe. L'onde se réfléchit au bout de la cochlée, et anime la chambre basse. Cela a pour effet de convertir les vibrations *longitudinales* de l'air en oscillations *transversales* (verticales, comme sur la figure 3, C à E) de la chambre médiane qui repose sur la membrane basilaire.

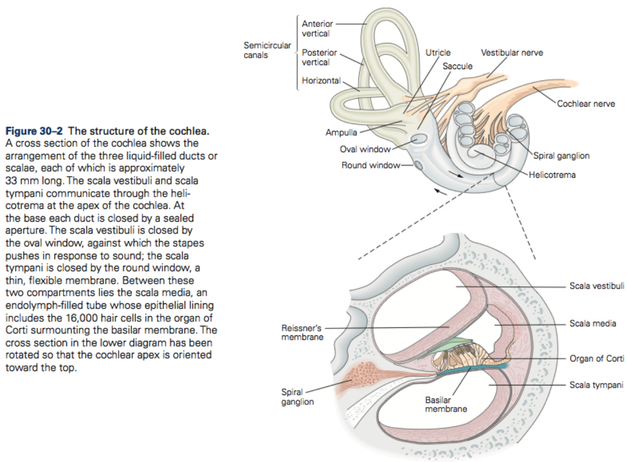


Figure 2 [31].

2. La chambre haute s'appelle *scala vestibuli*, et la chambre basse *scala tympani*.
3. Le terme technique est *apex* ou *helicotrema*.
4. Bien sûr, des neurones *arrivent* aussi à l'organe situé sur la membrane basilaire, pour effectuer les premières rétroactions : par exemple, la sensibilité des capteurs est ajustée en direct au niveau sonore global, et nous allons voir que les trois quarts des cellules-clé de la membrane ne sont pour ainsi dire présentes que pour le faire.
5. Marteau, enclume, étrier – *malleus, incus, stapes* de la figure 1.

Or, c'est dans cette chambre médiane que se trouvent les capteurs qui vont transmettre l'information sonore au reste du cerveau. Le premier fait important pour notre discussion est que la membrane basilaire a des propriétés physiques remarquables. Elle *s'élargit* à mesure qu'on s'éloigne de l'entrée de la cochlée (alors que la cavité s'amincit), et *sa rigidité change en chemin.* Une première conséquence de ces propriétés physiques fines est qu'un son *sinusoïdal* induit une vibration *bien localisée* de la membrane basilaire (figure 3D). Une deuxième est que les déformations induites par deux vibrations dont les fréquences sont différentes *ne sont pas localisées au même endroit* : la figure 3F résume l'effet d'une vibration sinusoïdale (d'un *ton pur*) en fonction de sa fréquence, et la figure 3G, l'effet d'un son comportant (comme il se doit !) plusieurs fréquences.

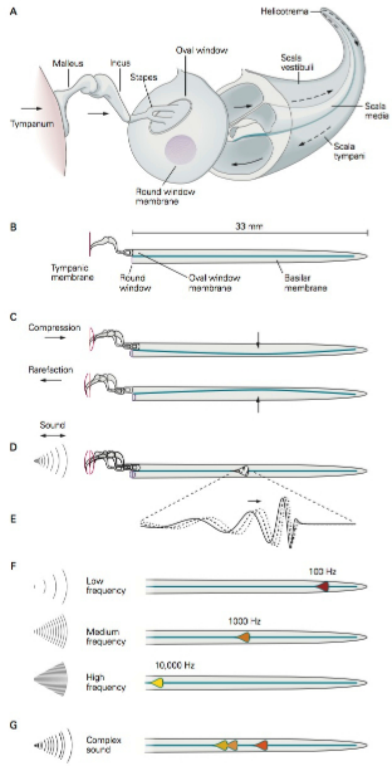


FIGURE 3 [31].

Les propriétés physiques de la membrane basilaire lui permettent ainsi de *séparer les fréquences*, et les nombreuses rétroactions sont là pour affiner ce rôle : la physique de la membrane est conditionnée par l'activité des cellules qui participent au traitement de l'information. La relation entre la fréquence d'un son sinusoïdal et la distance à l'entrée de la cochlée de la région qu'il active, qui est assez importante pour mériter le nom de relation *tonotopique*, est connue ; elle est... logarithmique !

La relation entre la fréquence d'un son et sa hauteur dans nos gammes musicales est logarithmique également. Jouez un accord sur votre clavier, et observez son effet sur la membrane basilaire : l'espacement des régions activées reproduit celui des touches que vous enfoncez [6].

Pour entendre, nous avons un clavier dans chaque oreille.

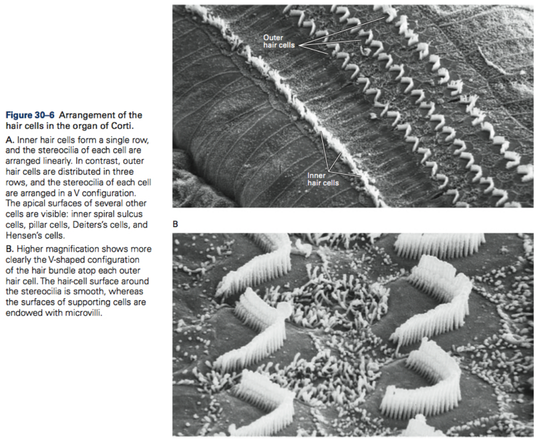
⋆

Venons-en aux capteurs eux-mêmes, ceux qui vont "activer" les neurones qui partent de la membrane basilaire pour former le premier nerf auditif. Ce sont les touches du clavier de notre oreille.

Leurs caractéristiques physiques sont bien sûr adaptées à la séparation des fréquences effectuée par la membrane, et nous allons voir qu'elles *renforcent* cette séparation.

6. Approximativement bien sûr, ne serait-ce que parce que nos claviers ordinaires ont choisi une gamme à sept tons et que la membrane basilaire n'a pas de touches noires. Mais deux intervalles identiques (deux octaves, deux quintes) activent bien deux couples de régions de même distance sur la membrane basilaire.

Observons une image de l'organe de Corti, qui est posé sur la membrane basilaire.



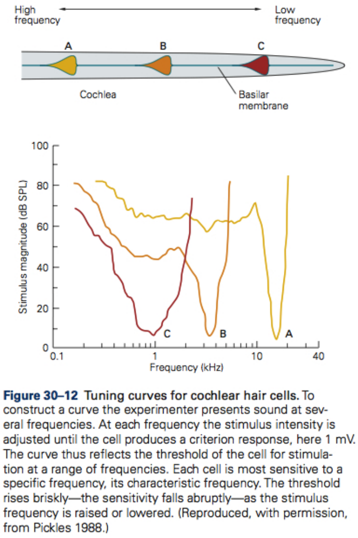


Figure 4 [31].

Au-dessus de chaque cellule se dresse un amas de "cils", et voici qu'apparaissent clairement quatre rangées parallèles de cellules qui signalent la présence de capteurs. Dans trois des quatres rangées, les amas de "cils" forment un beau chevron, et dans la quatrième, la plus importante, les amas sont approximativement en forme de pinceau.

C'est de la rangée intérieure, aux amas de cils en pinceaux, que part le signal. Pour que le taux de décharge des neurones situés en aval d'une de ces cellules soit important, il faut que l'oscillation de la membrane basilaire soit assez importante là où se trouve le capteur, et le taux de décharge induit par un ton pur (et par la déformation correspondante de la membrane) dépend de plus des propriétés physiques des cils, notamment de leur longueur. Chaque récepteur a ainsi une *fréquence préférée*, et une sensibilité plus ou moins grande aux fréquences très proches ; la figure de droite ci-dessus le résume.

Lorsqu'on stimule une cellule ciliée avec un ton pur dont la fréquence est sa fréquence préférée, le taux de décharge des neurones en aval dépend alors, dans la zone d'intérêt, à peu près linéairement [7] de l'intensité du son (mesurée bien sûr avec une échelle logarithmique de pressions – par exemple en décibels).

Les trois rangées extérieures, aux amas de cils en chevron, servent surtout à l'amplification du son (pour les discrets murmures) ; de façon remarquable, elles le font *en modifiant les propriétés physiques* de la membrane en fonction du signal courant. Cette rétroaction contribue grandement au pouvoir de discrimination de l'assemblée des cellules. Pour mesurer comme le système que je décris est délicat, il suffit de songer que le moindre désordre dans le bel alignement des rangées de chevrons peut signifier, même lorsque la perturbation est très petite, la surdité totale.

Nous avons environ seize mille cellules ciliées dans chaque oreille, et puisque la tonotopie est logarithmique, on peut utiliser le vocabulaire musical pour décrire la distance entre les fréquences préférées de deux capteurs adjacents. C'est : un trentième de demi-ton.

7. Je mets beaucoup de détails entre parenthèses ici : il y a par exemple des saturations, bien sûr ; le niveau global du son a une influence très importante *via* les trois rangées en chevron qui font l'objet de la remarque suivante, etc. Ces détails (non-linéarités statiques, influence du contexte et du niveau ambiant de signal) sont la règle en neurosciences.

Voici deux slogans qui résument grossièrement, mais qui résument, la situation [8] :

Ce que les capteurs de l'appareil auditif transmettent,
c'est la transformée de Fourier du son.

La "carte" indiquant comment les fréquences se répartissent le long de la cochlée
est connue : c'est un logarithme.

Pour qu'il soit clair que ces faits ne sont pas *que* mathématiquement séduisants, j'aimerais en rappeler ici une conséquence médicale. Lorsque la cochlée est endommagée, les *implants cochléaires* peuvent aller jusqu'à contourner [9] certaines formes de surdité profonde en enregistrant le son, en calculant sa transformée de Fourier, et en transmettant chaque coefficient de Fourier (par stimulation électrique) au neurone qui aurait dû le recevoir. Des dizaines de milliers de personnes en sont équipées en France.

⋆

Les deux slogans que je viens d'encadrer indiquent deux questions essentielles dans l'étude des structures nerveuses, et bien sûr des aires cérébrales, dédiées à une modalité sensorielle.

(a) **Spécialités des neurones :** comment *chaque* neurone réagit-il au stimulus ? Quelle information en extrait-il ? Ici, chaque neurone du nerf cochléaire extrait en première approximation un coefficient de Fourier du stimulus sonore (ou plutôt, naturellement, un coefficient de Fourier localisé sur une petite fenêtre temporelle, donc un coefficient de transformée en ondelettes) .

(b) **Cartes des spécialités :** comment les spécialités sont-elles *organisées géométriquement*, lorsqu'elles le sont, dans nos tissus, dans notre chair ? Ici, on trouve une carte des tons, dès la membrane basilaire sur laquelle se trouvent les rangées de capteurs.

Dans le cas de l'audition, nous venons de voir que des réponses à ces deux questions peuvent se lire *directement à l'endroit où se trouvent les capteurs*, et cela est tout à fait remarquable. Ce n'est pas le cas pour les autres modalités sensorielles, comme nous allons le voir. Les deux premières parties de mon travail tournent autour du fait que la théorie des groupes peut, je l'espère, nous aider à aborder certains aspects de ces deux questions dans d'autres contextes − pour le système visuel et pour le système vestibulaire. Les chapitres 1 et 6 de ma thèse s'occupent, chacun à sa manière, du comportement des neurones individuels, c'est-à-dire de la question **(a)**. Les chapitres 2, 3 et 4 partent de la façon dont les spécialités des neurones s'organisent − c'est la question **(b)** − dans les aires visuelles primaires. Quant aux chapitres 5, 7 et 8, bien que le chapitre 5 contienne des résultats qui préparent le chapitre 6, ils seront bien loin du cerveau.

8. Dans cette introduction, les encadrés bleus signalent des faits remarquables mais bien connus ; ce sont les encadrés verts de la section 3 qui isoleront les résultats principaux de ma thèse.

9. Note : ceci est l'introduction de ma thèse de mathématiques. Je n'ignore pas les questions éthiques autour de la surdité et je sais que le choix des mots compte, mais je rappelle que ma thèse ne porte pas sur le système auditif.

1.2 Profils récepteurs et cartes fonctionnelles dans le cortex visuel primaire

Les questions **(a)** et **(b)** ont reçu de célèbres éléments de réponse dans le cas du *cortex visuel primaire* des mammifères, plus précisément de sa première aire (aire 17 du chat) qui est l'aire visuelle la mieux connue du cerveau. Elle est située chez nous à l'arrière de la tête [10] ; c'est la première aire *corticale* [11] à recevoir des informations provenant de la rétine. Pour rappeler qu'elle intervient tôt dans le traitement de l'information visuelle, on abrège souvent son nom [12] en V1.

Pour ce qui suit, d'excellentes références sont les textes d'Hubel et Wiesel [28] (pour les premières découvertes), et le livre de Jean Petitot [46]. Le long texte de Daniel Bennequin [7] contient des discussions très adaptées à ce qui vient. Quatre chapitres de ma thèse partent d'études menées sur le cortex visuel primaire ; il est donc nécessaire que je donne quelques détails.

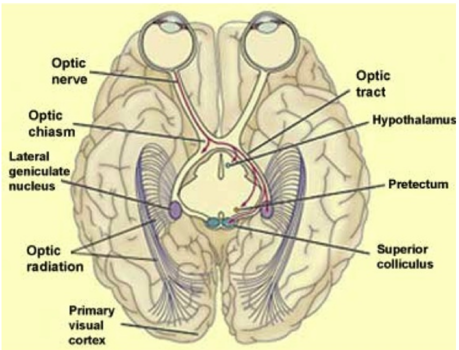


Figure 5.

La voie qui mène l'information visuelle de la rétine à V1 est représentée sur la figure ci-contre.

Les photorécepteurs de la rétine (bâtonnets et cônes) de chaque oeil transmettent (après plusieurs relais) l'information visuelle par l'activité électrique des neurones (les *cellules ganglionnaires*) qui forment le *nerf optique*. Chaque cellule ganglionnaire transmet l'information venant d'une centaine de photorécepteurs, et les photorécepteurs dont elle collecte les informations sont voisins [13] : chaque cellule ganglionnaire est donc reliée à un domaine précis de la rétine, c'est-à-dire du champ visuel [14].

Les deux nerfs optiques se rencontrent au *chiasme optique*, et leurs fibres recomposent deux faisceaux différents : parmi les fibres provenant de l'oeil gauche, une moitié − celle qui transmet les informations des bâtonnets et cônes qui ont reçu la lumière de la moitié *droite* du champ visuel − rejoint un ensemble des fibres provenant de l'oeil droit − celles qui transmettent les informations *de la même moitié du champ visuel*. Ensemble, ces fibres qui transmettent l'information de la moitié droite du champ visuel poursuivent vers l'hémisphère *gauche*, et le faisceau s'appelle désormais *tractus optique gauche*. Les autres fibres s'assemblent pour former le *tractus optique droit*, qui poursuit vers l'hémisphère droit.

10. Sa position n'est pas pour rien dans l'abondance de données sur le cortex visuel primaire : pour positionner des capteurs (hier des électrodes, aujourd'hui un système d'imagerie optique) près de V1, il "suffit" de pratiquer une ouverture à l'arrière de la tête...

11. Rappelons que *cortex* signifie *écorce* et que ce terme désigne la couche superficielle des hémisphères cérébraux, celle qui est plissée ; les hémisphères ont aussi des régions *sous-corticales* − ganglions de la base, hippocampe, amygdale... Par ailleurs, le système nerveux central ne se résume bien sûr pas aux hémisphères, et nous rencontrerons certaines de ses autres parties dans ce qui suit.

12. On parle aussi de "cortex strié" pour rappeler son anatomie ; ce terme est utilisé notamment par Hubel et Wiesel.

14. Ils occupent une surface de l'ordre du millimètre carré sur la rétine, dont la surface totale est de l'ordre du millier de millimètres carrés chez l'homme.

14. Jean Petitot dit : entre 0.5˚ et 10˚ d'angle visuel.

Par conséquent, chaque hémisphère reçoit des informations des deux yeux [15], mais ne reçoit des informations que sur la moitié contralatérale du champ visuel. Une lésion du *nerf* optique gauche entraîne une cécité partielle ou totale *de l'oeil gauche*, mais une lésion du *tractus* optique gauche entraîne une perte de vision *dans la moitié gauche du champ visuel* (une *hémianopsie*).

Cela dit, pour ce que je vais présenter mainenant il n'y a pas d'autre différence entre les hémisphères cérébraux : dans ce qui suit, je ne les distinguerai donc plus et je parlerai par exemple de *l'aire V1*, au singulier, plutôt que de distinguer les aires des deux hémisphères.

Pour l'instant, nous n'avons suivi qu'un axone depuis la rétine. Les voies visuelles sont diverses (il y en a sept ou huit), et certains neurones vont à l'hypothalamus pour participer à des fonctions de régulation comme celle du cycle veille/sommeil, d'autres projettent dans le colliculus qui coordonne les mouvements des yeux. Mais plus de 90% des cellules du nerf optique projettent directement dans le thalamus, dans le *corps genouillé latéral.* Je reviendrai rapidement sur les neurones du corps genouillé latéral au début du chapitre 1 ; pour que cette introduction déjà longue ne devienne pas démesurée, je prie ma lectrice ou mon lecteur de s'y reporter. En effet, immédiatement après les synapses qui suivent le corps genouillé latéral, au bout des axones qui forment la *radiation optique* de la figure 5, nous arrivons dans V1. Les précisions que j'ai données sont suffisantes pour que nous puissions revenir aux questions **(a)** et **(b)** ci-dessus lorsqu'elles concernent V1.

Chez l'homme, l'aire V1 rassemble plus de cent millions de neurones [16]. Le nerf optique, à travers lequel la totalité de l'information visuelle transite, rassemble un bon million de neurones. Chaque cellule du nerf optique projette donc sur de nombreux neurones de V1 ; réciproquement, un neurone de V1 reçoit des signaux de plusieurs neurones du corps genouillé latéral ; nous allons voir comment cela lui permet de raffiner sa spécialisation.

Signalons, pour compléter notre présentation sommaire, qu'il y a de très nombreuses rétroactions entre le corps genouillé latéral et V1, et que les retours venus de V1 ont un rôle important dans la modulation de l'activité des cellules du corps genouillé latéral et d'autres noyaux thalamiques ; par ailleurs, V1 reçoit des retours importants de plusieurs aires visuelles ultérieures.

(a) Les caractéristiques des neurones de V1. La notion de profil récepteur.

Des éléments de réponse à la question **(a)** ont été rendus célèbres par David Hubel et Torsten Wiesel. Le récit de leurs découvertes est tout à fait passionnant, et je renvoie au discours de réception du Prix Nobel [29].

Je vais parler surtout des *cellules simples* de V1, celles qui sont directement reliées aux cellules du corps genouillé latéral ; la plupart des neurones de V1 reçoivent eux-mêmes des informations de plusieurs cellules simples et leurs caractéristiques sont héritées pour partie de celles des cellules simples, de sorte que beaucoup des caractéristiques dont je vais parler concernent une proportion bien plus importante de neurones de V1. De façon générale, dans V1,

15. C'est très utile : chaque hémisphère peut ainsi comparer les informations des deux yeux. C'est par exemple un des supports de la perception de la profondeur et de l'impression de relief, la stéréoscopie et les "lunettes 3D" ne s'y trompent pas.
16. Les estimations varient entre 100 et 500 millions.

– chaque neurone a une *position préférée* : il est sensible à la structure de la scène visuelle dans un domaine restreint du plan visuel (dont la taille dépend du neurone et du contexte). La structure de la scène visuelle hors de ce domaine a peu d'influence sur l'activité électrique du neurone [17].
– chaque neurone a une *orientation préférée* : il réagit préférentiellement à des stimuli allongés dans une direction précise (lorsqu'ils sont au moins en léger mouvement), et lorsqu'un stimulus allongé est présenté dans le domaine de la scène visuelle qu'il observe, l'activité électrique du neurone décroît à mesure que l'orientation du stimulus s'écarte de celle qui maximise la réponse du neurone, comme sur la figure 6.
– chaque neurone a une *fréquence spatiale préférée* : lorsqu'on présente, dans sa région préférée, une "grille" de barres d'espacement régulier dont l'orientation est son orientation préférée, l'activité électrique du neurone dépend de l'espacement des barres ; elle décroît lorsque l'espacement s'éloigne de l'espacement préféré du neurone.

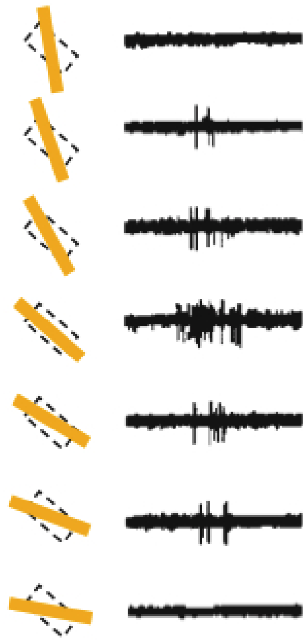


Figure 6.

Il y a d'autres spécialités, bien sûr : oeil préféré (ou degré de binocularité) lorsque la position préférée est dans la région vue par les deux yeux (c'est la notion de *dominance oculaire* des neurones de V1) , couleurs, vitesse des stimuli, sens du mouvement, fréquence temporelle... Je ne m'y intéresserai pas ici.

Les trois spécialités sur lesquelles je viens d'insister peuvent être rassemblées dans la notion de *profil récepteur*, dont voici un résumé (voir notamment le chapitre 2 du manuel de Dayan et Abbott [15]).

Oublions un instant que nous voyons en couleurs, et résumons la scène visuelle à des niveaux de gris se détachant sur un plan. Cela revient à assimiler l'image qui vient à la rétine à l'instant t à une fonction, disons $\mathcal{I}_t$, de $\mathbb{R}^2$ dans (un intervalle de) $\mathbb{R}$. Nous sommes en biologie, et il n'y a pas de mal à supposer que le support de $\mathcal{I}_t$ est compact et, bien que notre rétine ait une résolution finie, que $\mathcal{I}_t$ est continue.

L'image $\mathcal{I}_t$ est bien sûr essentielle pour déterminer l'activité électrique d'un neurone de V1, sinon à t, du moins peu après t (même lorsqu'elle est modulée par les nombreuses rétroactions). Voyons quel peut être son effet. Observons une *cellule simple* qui analyse les caractéristiques locales de l'image autour de $\mathbf{x}_0$, dont l'orientation préférée pour les stimuli visuels est ϑ_0, et la fréquence préférée, κ_0.

17. Elle en a, naturellement, mais à travers des rétroactions d'autres neurones corticaux : il suffit de se rappeler l'exemple que je donnais d'un même lieu vu de jour et de nuit et reconnu sans peine, alors que les luminosités sont quantitativement presque incomparables, pour voir que le contexte a une influence importante. Cette sensibilité à la luminosité globale, qui s'accompagne une sensibilité analogue au contraste, existe dès la rétine.

Choisissons l'un des deux vecteurs de norme κ_0 et de direction ϑ_0, notons-le $\mathbf{k_0}$ et partons d'un *filtre de Gabor* associé à ces données, c'est-à-dire de la fonction

$$\mathbf{x} \mapsto \exp\left(\frac{\|\mathbf{x}-\mathbf{x}_0\|^2}{2\sigma^2}\right)\cos\left(\langle \mathbf{k_0}, \mathbf{x}\rangle + \varphi\right)$$

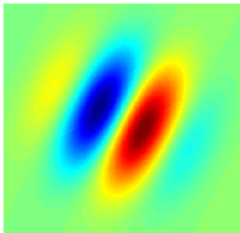


FIGURE 7.

où σ est un réel positif que nous allons pouvoir interpréter dans un instant comme déterminant la taille de la région du plan visuel analysée par le neurone, et où la phase φ détermine quelles sont les régions inhibitrices et excitatrices.

Disons qu'une fonction $\mathcal{P} : \mathbb{R}^2 \to \mathbb{R}$ est un *profil récepteur* d'un neurone visuel lorsque le produit scalaire

$$\langle \mathcal{P}, \mathcal{I}_t \rangle_{\mathbb{L}^2}$$

permet d'obtenir une description raisonnable de l'activité électrique du neurone que nous observons, par exemple en lui appliquant une fonction logistique (sigmoïde) pour tenir compte des habituelles saturations et en adoptant le résultat pour une mesure du taux d'émission de potentiels d'action à l'instant t ou à un instant ultérieur [18].

Nous pouvons alors résumer par le slogan suivant les précisions apportées, depuis les années 1970, aux résultats d'Hubel et Wiesel.

Les filtres de Gabor fournissent de bons modèles pour les profils récepteurs des cellules simples de V1.

Autrement dit :

Ce qu'un neurone du cortex visuel primaire extrait de l'image, c'est un coefficient de transformée en ondelettes.

Compte tenu des propriétés des ondelettes pour le traitement de l'information (et bien sûr de leur rôle pour la compression d'images), c'est tout à fait remarquable.

Au chapitre 1, je reviendrai rapidement sur l'explication proposée par Hubel et Wiesel pour la construction de ce profil récepteur à partir de ceux des neurones du corps genouillé latéral afférents à un neurone donné de V1. Mais pour faire voir d'où vient l'essentiel de la première partie qu'on va lire, il faut maintenant que je revienne à la question **(b)** ci-dessus, et que je dise un mot de la façon dont les spécialités des neurones sont disposées géométriquement dans le cortex (et donc comment l'assemblée des neurones se répartit les divers coefficients de Fourier locaux de l'image).

(b) Les cartes fonctionnelles du cortex visuel primaire

Commençons par dire que le cortex visuel, puisque c'est un morceau de cerveau, a trois dimensions. Mais le cortex [19] n'est épais que de quelques millimètres, et on sait depuis Ramon y Cajal (au dix-neuvième siècle) qu'il est organisé en fines *couches* parallèles à la surface (disons extérieure).

Pour répondre à la question **(b)**, les méthodes expérimentales ont longtemps fait usage d'une unique électrode qu'on déplaçait de neurone en neurone, pour voir comment les spécialités changeaient. Aujourd'hui, l'imagerie optique permet de les voir "de haut" dans une

18. Dans de nombreux contextes, cette notion figée ne suffit pas : il faudrait tenir compte de l'évolution de l'image et introduire un profil *spatiotemporel*, fonction sur $\mathbb{R}^2 \times \mathbb{R}$, et en prendre le produit scalaire $\mathbf{L}^2$ avec $(x,t) \mapsto \mathcal{I}_t(x)$ Pour pouvoir envisager la collaboration entre le système visuel et le système vestibulaire, il faudrait aussi élargir ces profils spatiotemporels à (2+1) dimensions en profils à (3+1) dimensions.

19. *Écorce*, comme je l'ai dit plus haut.

fenêtre plus large, projetées depuis la surface du cortex sur celle du système d'enregistrement. Cela dit, pour qu'une telle information soit pertinente, il faut avoir un renseignement sur l'organisation dans la direction transverse. Dès la fin des années 1950, le fait suivant était clair.

Lorsqu'on s'enfonce dans la direction normale à la surface du cortex, aucune des trois spécialités sur lesquelles j'ai insisté ne varie.

Cela ne signifie pas que le profil récepteur ne change pas : par exemple, la phase change, la taille de la fenêtre qu'a le neurone sur le monde (le paramètre σ si le profil récepteur est décrit par un filtre de Gabor) change, sa précision dans la discrimination des orientations n'est pas constante même si l'orientation préférée l'est, sa sensibilité au contraste varie également... mais la position préférée du neurone, son orientation préférée et sa fréquence spatiale préférée ne changent pas pour l'essentiel.

Par conséquent, lorsqu'on s'intéresse à la carte des spécialités des neurones, on peut raisonnablement résumer le cortex à un morceau $\mathcal{C}$ d'une surface plongée dans $\mathbb{R}^3$. Cette surface n'est bien sûr pas plate, et cela jouera un rôle dans le chapitre 3 ; mais elle est molle, et on peut tout à fait l'imaginer dépliée sur un morceau de plan. Compte tenu du nombre de neurones dans V1, on oubliera bien sûr le fait qu'un neurone occupe un espace fini et on imaginera qu'il y a un neurone en chaque point de $\mathcal{C}$ (ou plutôt une *microcolonne*, rassemblant tous les neurones situés "sous" ce point).

(b,1) La notion de carte neurale ; la rétinotopie.

Parlons de la carte des positions préférées des neurones ; puisque chaque neurone a une position préférée, il y a une application *rétinotopique*[20] de $\mathcal{C}$ vers le plan visuel $\mathcal{V}$. Cette application est continue : deux neurones proches ont des positions préférées proches. C'est un homéomorphisme.

Mais ce n'est pas une isométrie, ni une transformation affine. Cela n'a rien de surprenant, c'est un fait connu du grand public pour d'autres cartes neurales : l'application *somatotopique* de la surface de la peau vers celle du cortex somatosensoriel est aussi (par morceaux) un homéomorphisme, mais la majorité de la surface du cortex somatosensoriel s'occupe de petites parties du corps (les doigts, les lèvres, la langue...) ; "l'homoncule" auquel elle donne lieu lorsqu'on compose son inverse avec un difféomorphisme de $\mathbb{R}^3$ qui en fait une isométrie est célèbre.

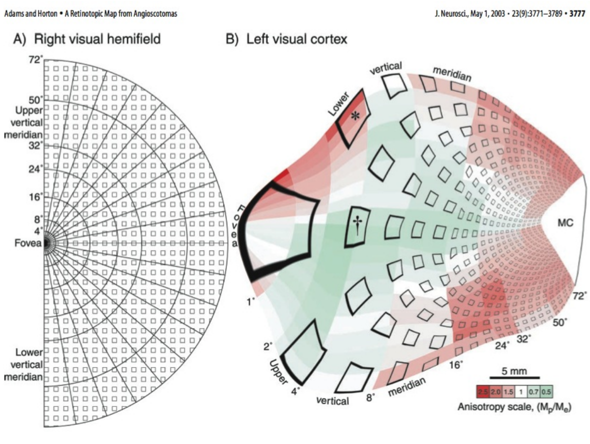


Figure 8.

De nombreuses données sont disponibles sur l'application rétinotopique de $\mathcal{C}$ vers une partie (une moitié) du plan visuel $\mathcal{V}$: par exemple, il est habituel de dire qu'elle est conforme (et même logarithmique) chez le chat et le singe. Elle n'interviendra pas directement dans ce qui suit, mais en avoir une idée sera utile pour les motivations du chapitre 3 : la figure 8, due à Adams et Horton, donne une représentation assez précise de sa réciproque – la figure a été obtenue en étudiant la propagation des angioscotomes chez le singe[21].

20. On pourrait dire, et on dit parfois : visuotopique.

21. Les angioscotomes sont des "taches aveugles" dans le champ de vision dont l'origine est vasculaire.

(b,2) Cartes d'orientation, pinwheels.

La carte des positions préférées est un homéomorphisme conforme entre une partie de $\mathcal{C}$ et $\mathcal{V}$ comme nous venons de le voir. Celle des orientations préférées, puisque c'est une application de $\mathcal{C}$ dans $\mathbf{P}^1(\mathbb{R})$, n'en est pas un et possède de la redondance. Si la carte des orientations est plus célèbre encore que celle des positions, c'est qu'on sait depuis Hubel et Wiesel que cette redondance est essentielle pour la perception : le but de l'arrangement des spécialités d'orientation semble être que la zone de V1 qui analyse un petit domaine donné du plan visuel ait accès à toutes les orientations, de façon à répartir sur le plan cortical l'espace $\mathcal{V} \times \mathbf{P}^1$ es directions locales, qui est de dimension 3. Commençons par énoncer sous forme de slogan la raison pour laquelle la géométrie des cartes d'orientations est célèbre avant d'expliquer ce que le slogan recouvre.

*La "carte" indiquant comment, à la surface du cortex, se répartissent
les spécialités d'orientation,
a des propriétés géométriques remarquables et communes à de nombreuses espèces.*

Le point de départ des chapitres 2 à 4 de mon texte, c'est le rôle de l'invariance (par une action de groupe) dans les modèles théoriques qui ont été proposés pour essayer de comprendre ces propriétés remarquables. Il faut donc que je donne quelques détails (voir le chapitre 2 et le début du chapitre 3).

Les cartes d'orientation sont maintenant obtenues par imagerie optique — à côté de techniques d'imagerie biphoton, assez fines pour obtenir des détails *au neurone près* (voir [45]). Une manière traditionnelle de les représenter est d'assigner une couleur à chaque orientation, et de colorier chaque point de $\mathcal{C}$ avec la couleur qui représente l'orientation préférée commune aux neurones situés "sous" ce point. La figure ci-dessous, due à William Bosking et ses collègues, montre une carte d'orientation typique (voir le début du chapitre 3 pour le protocole expérimental qui permet d'obtenir une telle figure).

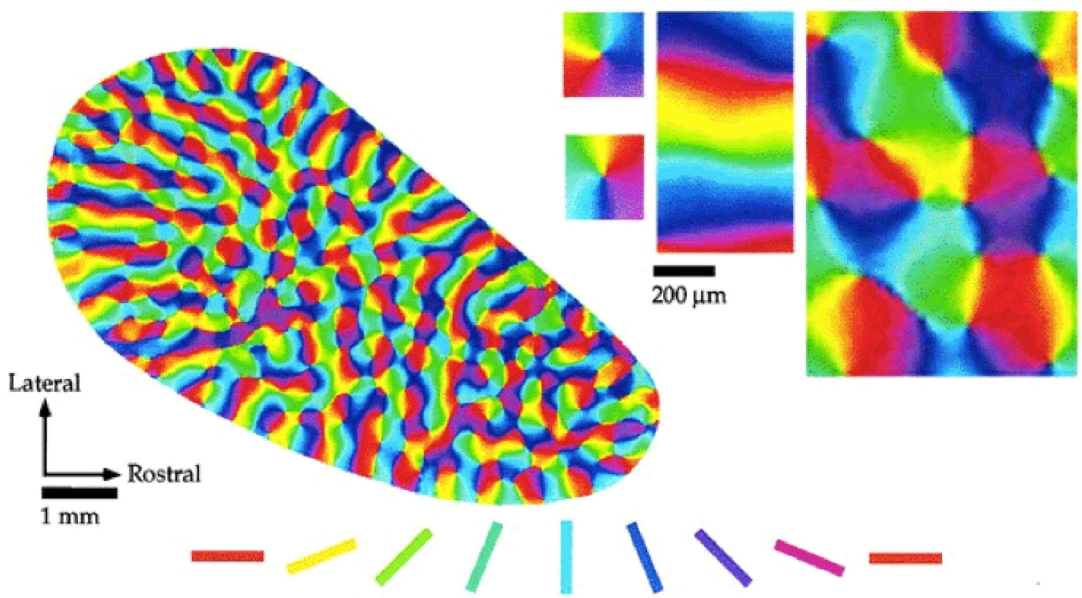


FIGURE 9 [11].

Je commenterai abondamment cette carte aux chapitres 2 et 3. Pour cette introduction, relevons-en deux traits essentiels :

La carte des orientations préférées est, en un sens vague, globalement quasipériodique, au moins dans la région centrale : elle a une "longueur d'onde" caractéristique, disons Λ, de sorte que deux domaines dont les centres sont à distance Λ ont une grande probabilité d'avoir un coloriage proche.

⋆

*Elle présente des **singularités** : il y a des points qui, dans chacun de leurs voisinages, voient toutes les orientations représentées. Ce sont les centres des **pinwheels**.*

Je serai amené à revenir plusieurs fois sur le premier point, alors disons rapidement pourquoi la "quasipériodicité" que je mentionnais est importante pour le cerveau. Hubel et Wiesel ont constaté que lorsqu'on parcourt un petit segment (de longueur inférieure à, disons, 1mm) sur la surface corticale, les préférences d'orientation varient assez vite pour que toutes les orientations soient représentées sur le segment, alors que la position préférée des neurones varie peu (quelques degrés d'angle visuel tout au plus). De plus, dans un domaine de taille adéquate sur le plan cortical (par exemple si Λ est la longueur caractéristique de la carte, de sorte qu'il y a une grande probabilité de rencontrer toutes les couleurs sur un segment de longueur Λ, dans un carré de côté Λ^2), les neurones analysent une région restreinte du plan visuel, mais les autres spécialités y sont généralement représentées de façon surjective : on trouve des neurones analysant toutes les directions, une large gamme de fréquences spatiales, tous les degrés de binocularité, toutes les couleurs...

Hubel et Wiesel ont suggéré que cela permettait à V1 d'analyser le stimulus visuel en faisant jouer à ces domaines le rôle d'*unité fonctionnelle* analysant toutes les caractéristiques locales de l'image dans une petite région du plan visuel. Ils ont proposé le nom d'*hypercolonne* pour ces unités fonctionnelles [22].

Notons cependant que cette suggestion ne revient pas à dire qu'il existe une *partition* de la surface du cortex en unités distinctes que l'on pourrait identifier comme des hypercolonnes et dessiner sur la carte, mais plutôt que chaque région de taille adéquate (Λ^2) est susceptible d'avoir un rôle fonctionnel important. La notion d'hypercolonne est très importante pour notre compréhension du fonctionnement de V1, mais ses contours mathématiques, bien qu'il soit clair que la quasipériodicité de la carte y joue un rôle, sont assez vagues [23]. Ce fait ne sera pas sans influence dans les chapitres 2 à 4.

Je vais maintenant décrire un résultat expérimental remarquable obtenu récemment par le groupe de Matthias Kaschube et Fred Wolf à l'Institut Max Planck de Göttingen. C'est ce résultat, et le rôle que semblent devoir jouer les arguments de symétrie si on souhaite le comprendre abstraitement, qui a motivé les chapitres 2 à 4 de ma thèse. Il porte sur la *densité des pinwheels* [24], c'est-à-dire sur le nombre moyen de centres de pinwheels dans

22. Plus généralement, l'idée que les neurones corticaux peuvent se grouper en *modules fonctionnels*, auxquels la connectivité cortico-corticale permet de collaborer, joue depuis un grand rôle en neurosciences.

23. Les usages de la notion sont d'ailleurs assez divers et assez peu compatibles pour qu'il ne soit pas sûr qu'une définition formelle soit souhaitable.

24. Pour comprendre d'où vient le terme "pinwheel", il suffit de le taper dans un moteur de recherche d'images : dans le langage courant, c'est un jouet dont la forme et les couleurs sont évocateurs.

une région dont l'aire est le carré Λ^2 de la longueur d'onde caractéristique de la carte.

Il y a au moins deux possibilités pour définir une distance Λ qui permette de parler quantitativement de la longueur d'onde caractéristique.

– Une première méthode est locale : en chaque point, on obtient une estimation de longueur caractéristique locale − en notant d'abord, sur chaque demi-droite, la distance à laquelle se situe le point le plus proche ayant la même spécialité d'orientation, puis en prenant la moyenne sur les directions des demi-droites. On prend enfin pour Λ la moyenne des résultats locaux.
– Une deuxième est tout à fait globale : on part de la carte donnée, et on forme la fonction de $\mathcal{C}$ dans $\mathbb{R}$ qui donne son coefficient de corrélation avec chacune de ses translatées ; la transformée de Fourier de cette fonction est avec une bonne précision radiale, la façon dont son module dépend de la distance à l'origine présente un unique pic, et on choisit pour Λ la longueur d'onde qui indique la position du sommet de ce pic spectral.

Les deux méthodes donnent des résultats très proches [25] sur les cartes réelles (bien que ce ne soit pas le cas, comme nous allons le voir au chapitre 2, pour des candidats mathématiques naturels à reproduire ces cartes).

Kaschube, Schnabel, Lowel, Coppola et Wolf ont mesuré la densité des pinwheels chez des mammifères aux genres de vie très différents (le galago est un petit primate qui vit la nuit au creux des arbres et mange de tout, le furet tient compagnie le jour aux familles d'Europe et d'Amérique et est exclusivement carnivore, le tupaya est actif jour et nuit entre le sol des forêts et leurs arbres), et issus de lignées évolutives séparées depuis longtemps. Voici leur étonnante conclusion.

Au sein d'une espèce, l'écart entre les densités de pinwheels des cartes individuelles est faible (entre 2.5 et 3%) ; la moyenne pour chaque espèce est étrangement proche [26] de 3.14.

C'est assez étonnant pour mériter un slogan :

La densité de singularités observée sur les cartes du cortex visuel des mammifères, c'est π.

Si Kaschube et ses collègues se sont intéressés à la densité des pinwheels chez des animaux très différents, c'est que les conditions du développement des cartes corticales faisaient débat : la belle et efficace géométrie des cartes d'orientation est-elle programmée génétiquement, ou apparaît-elle par auto-organisation pour optimiser les performances de V1 dans les tâches visuelles, à mesure que le jeune animal acquiert de l'expérience visuelle ? Il y a longtemps que Wolf et Geisel prédisaient qu'une densité de π serait la signature d'un développement auto-organisé plutôt que programmé génétiquement. Pour soutenir cette idée ils avaient proposé un modèle de développement précis, sur lequel je reviendrai rapidement au chapitre 3.

25. Les méthodes que je viens de décrire correspondent aux définitions que j'adopterai aux chapitres 2, 3 et 4 pour démontrer des résultats de mathématiques, mais celles qui sont utilisées en pratique sont bien sûr plus subtiles (voir [33], p. 4). La méthode locale utilisée par Kaschube et ses collègues n'est pas celle que j'ai décrite : elle consiste à chercher quelle est la longueur d'onde qui maximise la moyenne, prise sur toutes les directions de propagation, des coefficients de transformée en ondelettes de Gabor qui sont localisés près de ce point. La méthode globale est plus proche de celle que j'ai décrite, mais sa mise en oeuvre est délicate et elle identifie le maximum en ajustant ("fit") les paramètres d'une expression *ad hoc* pour le spectre de puissance.

26. 3.12 pour le tupaya, 3.15 pour le furet, 3.15 pour le galago

Les chapitres 2, 3 et 4 de ma thèse reprennent et généralisent les faits mathématiques qu'ils ont découverts en chemin et qui les ont menés à prédire la densité π.

J'expliquerai en détail pourquoi ces faits sont *conséquence directe d'hypothèses de symétries dans les modèles.* Les modèles théoriques qui essaient de décrire la structure des cartes du cortex visuel primaire, notamment la densité π, utilisent *tous* une action du groupe des déplacements sur l'ensemble des cartes corticales admissibles ; il sera bientôt temps d'en parler.

⋆

Avant de quitter les cartes corticales pour quelques pages, disons un dernier mot sur la répartition des spécialités des neurones du cortex visuel primaire. Nous venons de voir que des données précises et robustes existent pour la carte des positions préférées (la rétinotopie) et pour la carte des orientations préférées. Pour la carte des fréquences spatiales, rien de tel en revanche (voir [51]) : le débat est brûlant d'actualité, même sur leur existence, et il y a peu de données ; de modèles, moins encore (voir cependant [52]).

1.3 Le système vestibulaire. Le cervelet vestibulaire.

Je viens de rappeler pourquoi il n'est pas absurde de dire que la cochlée et le cortex visuel primaire appréhendent, à chaque instant, l'espace des sons et l'espace des images grâce à des transformées de Fourier localisées.

Je vais maintenant présenter le fonctionnement des capteurs liés à une modalité sensorielle parfois méconnue, mais cruciale pour notre insertion dans notre environnement et probablement pour notre conception du monde : la perception du mouvement. Pour ce qui suit, d'excellentes références sont Berthoz [10], Angelaki et Cullen [3].

Revenons dans l'oreille ; je vais expliquer ce que sont les capteurs du système vestibulaire.

Le labyrinthe osseux de notre oreille n'abrite pas que la cochlée : elle en est la partie inférieure, et au-dessus de la partie médiane (qu'on appelle *vestibule* parce que vue du tympan, elle semble servir d'entrée au labyrinthe), on trouve trois boucles osseuses (figure 10). Ces boucles protègent des canaux, les *canaux semi-circulaires*, qui sont des membranes remplies du même liquide que la cochlée (l'endolymphe). Le vestibule comporte quant à lui deux poches dans lesquelles sont situés deux organes dont la surface semble parsemée de cailloux, utricule et saccule sur les figures 2 et 11 : ce sont les *otolithes*.

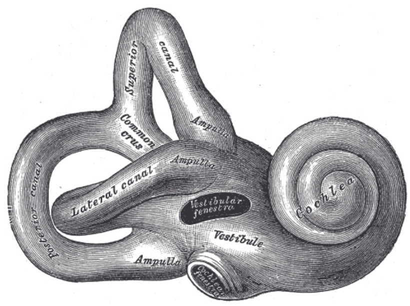


FIGURE 10 [22].

Galien avait déjà remarqué l'existence de ces structures (au moins des canaux), sans parvenir à préciser leur fonction : pendant deux mille ans, on a pensé que l'oreille ne servait qu'à entendre.

À la fin du dix-neuvième siècle, Darwin attribue encore le vertige à des causes digestives ou musculaires. Mais vers 1820, Purkinje met cette idée à mal en observant des patients placés sur une chaise tournante ; il pense que le vertige est le résultat d'un conflit entre la vision et la perception du mouvement.

À peu près à la même époque, Flourens sectionne les canaux semi-circulaires d'un pigeon ; l'animal semble toujours entendre, mais se met à marcher d'une bien étrange façon : il perd l'équilibre, se cache dans des coins obscurs. Plus étrange encore, lorsqu'un seul des trois canaux est sectionné, l'animal se met à "tourner en rond" ou (selon le canal sectionné) à confondre le haut et le bas. Mais Flourens n'en tire pas de conclusion claire.

La fonction des canaux semi-circulaires est précisée brusquement en 1873 : deux savants viennois, aujourd'hui célèbres pour d'autres recherches, la mettent au jour indépendamment et simultanément.

Le premier est physicien, c'est Ernst Mach. Après avoir étudié les mouvements de rotation pour la physique, il se rend compte qu'il *doit* exister un organe fait pour percevoir les mouvements de rotation de la tête. Il est au courant des résultats de Purkinje, et désigne les canaux semi-circulaires comme l'organe idéal pour remplir cette fonction.

Le second est médecin, c'est Josef Breuer. Il reproduit les expériences de Flourens, mais plus méthodiquement, et formule la même conclusion que Mach. Nous allons voir comment ses résultats permettent de la préciser.

Canaux semi-circulaires et otolithes sont notre centrale inertielle :
c'est avec eux que nous enregistrons les mouvements de notre tête.

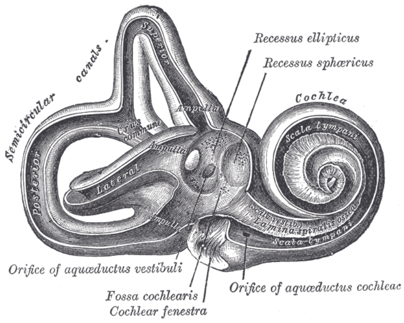


Figure 11 [22].

La façon dont les canaux semi-circulaires détectent les rotations a été détaillée par Breuer. Lorsqu'un canal subit une accélération de rotation, l'endolymphe qui y séjourne n'est pas entraîné tout de suite ; l'ampoule (*ampulla*) située au bout du canal voit la pression du liquide changer.

Or, c'est au bout de l'ampulla que sont situés les capteurs d'où partent les neurones qui transmettent l'information. Sa constitution physique est tout à fait remarquable (je ne vais pas détailler ici, voir [14]), et permet de transmettre de nombreuses informations, par exemple l'accélération angulaire autour de l'axe orthogonal au plan du canal, notamment pour les mouvements dont les fréquences sont inhabituellement hautes.

Cela dit, Harline, Terzuolo et d'autres ont étudié finement la dynamique des capteurs de l'ampulla dans les années 1960, et conclu que pour les mouvements associés aux déplacements "ordinaires" de l'animal [27], les neurones situés en aval d'un canaux transmettent surtout *la vitesse angulaire de rotation autour d'un axe privilégié* (chez l'homme où les canaux sont presque plans, l'axe orthogonal au plan du canal.)

Quant aux cristaux des otolithes, ils sont posés sur une base à peu près [28] plane (la base s'appelle *macula* ici), et ils sont isolés de l'endolymphe par une membrane. Lorsque la tête subit une accélération *linéaire* parallèle à la macula, les cristaux subissent une force en sens contraire. Or, sous les cristaux est situé un amas de cellules ciliées (comme sur la membrane basilaire) reliées des neurones ; l'activité électrique des neurones situés en aval de chaque cristal est alors (en gros, et dans la zone d'intérêt) *proportionnelle à l'accélération linéaire* selon un axe qui dépend de la cellule ciliée.

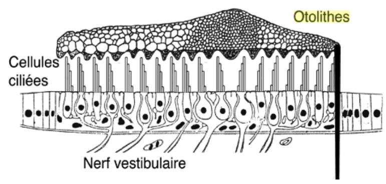


Figure 12 [10].

Les canaux semi-circulaires enregistrent les vitesses de rotation de la tête, les otolithes enregistrent ses accélérations linéaires.

⋆

Le système vestibulaire doit nous renseigner sur l'état de mouvement de notre tête. Or, il est clair que les sorties des capteurs ne sont pas directement suffisantes pour le faire :
– Pour avoir une idée de l'orientation de notre tête et de sa vitesse, il faut *intégrer* ces informations instantanées sorties des capteurs.

27. Les fréquences des mouvements ordinaires d'un animal dépendent de sa taille : ce ne sont pas les mêmes pour la souris et pour l'éléphant ! Pour cette raison, la géométrie des canaux est finement adaptée à chaque espèce.

28. Elle n'est pas *vraiment* plane, et c'est probablement important pour le traitement de l'information vestibulaire : pour une étude récente, voir [16].

– Un mouvement rectiligne et uniforme n'a bien sûr aucune répercussion sur les canaux ni sur les otolithes.
– Du point de vue des otolithes, la contribution de la gravité introduit une ambiguïté dans le signal d'accélération linéaire. D'ailleurs, une simple *inclinaison statique* de la tête, en créant une composante non nulle de la gravité dans le plan de la macula, induit une force constante sur les cellules ciliées, et sans travail de traitement de l'information, cette force peut être confondue avec celle qui résulterait d'une accélération de la tête dans le plan horizontal.

Pour le premier problème, il est essentiel pour nous de noter que l'inclinaison de la tête ne s'obtient pas à partir des vitesses angulaires par une intégration "simple" : c'est *dans le groupe des rotations* qu'il faut intégrer une équation différentielle pour obtenir une information globale à partir des informations instantanées captées par les canaux semi-circulaires.

Les deux autres problèmes sont si connus de la physique, et le cadre conceptuel pour en parler est si nettement proche du sujet de cette thèse et adapté au premier problème, qu'il est peut-être utile d'énoncer à nouveau l'information transmise par les canaux et otolithes à l'aide du slogan suivant (il sera précisé aux chapitres 5 et 6).

Ce qui est transmis par les canaux semicirculaires et par les otolithes, c'est un élément de l'algèbre de Lie du groupe de Galilée homogène.

⋆

Comment les structures cérébrales qui traitent l'information vestibulaire réussissent-elles à tirer de ces informations infinitésimales un renseignement global sur notre mouvement, en séparant au passage la contribution de la gravité de l'accélération linéaire due à nos mouvements ? Où l'intégration dans le groupe des rotations se produit-elle dans le cerveau ?

Compte tenu du rôle de l'analyse harmonique dans les réponses à la question **(a)** de la section 1.1 lorsqu'elle concerne l'audition ou la vision, et compte tenu des rapports de l'analyse de Fourier classique avec la structure du groupe des translations (voir la suite de cette introduction), y a-t-il des neurones qui gèrent le stimulus vestibulaire *en utilisant la structure du groupe de Galilée* ? La collaboration entre le système visuel et le système vestibulaire, essentielle à bien des titres (notamment pour lever l'ambiguïté induite par l'absence d'effet des mouvements rectilignes uniformes sur les capteurs vestibulaires) et mal comprise à l'heure qu'il est, serait-elle plus abordable en utilisant la structure du groupe de Galilée *inhomogène* pour mêler les translations de l'espace-temps (avec trois dimensions d'espace) essentielles pour la vision et le groupe de Galilée homogène de l'information vestibulaires ?

Si ces questions admettent des formulations mathématiques précises, on y trouvera probablement la théorie des groupes et l'analyse harmonique invariante.

1.4 Des groupes ?

Je rappelais en ouvrant cette introduction que nous devons nous adapter au fait que notre environnement change, et que nous bougeons par rapport à lui. Un objet familier ou dangereux doit être reconnu vite, même lorsque nous sommes en pleine activité, et nous devons adapter nos mouvements et nos perceptions à ses changements réels ou apparents.

C'est une bonne raison pour laquelle les groupes doivent être d'un précieux secours pour décrire ce qui, dans nos structures cérébrales, sous-tend notre rapport à l'espace et au mouvement. La suggestion est très ancienne, et remonte à Helmholz [25] et Poincaré [49] (voir aussi Cassirer [12]). Certains n'hésitent pas à affirmer que les groupes sont un support (explicite ou implicite) des opérations mentales (ou neuronales, faut-il chercher à séparer ?) qui fondent notre rapport à l'espace ; Piaget [48] est l'un des plus clairs.

La structure des groupes de déplacements euclidiens (plans ou spatiaux) semble jouer un grand rôle. Pour comprendre pourquoi, rappelons (bien que le livre qu'elle ouvre ne parle pas du cerveau) la belle image de Souriau [58] :

> *Devant vous, un écran s'allume. Au centre, vous apercevez la lettre* **S** *; d'autres lettres sont réparties sur l'écran, inclinées dans tous les sens ; certaines retournées, d'autres pas. Le jeu est simple : la même lettre* **S** *figure une seconde fois quelque part sur l'écran, il faut réussir à l'y pointer avec une commande. Le plus vite possible : vous serez chronométré.*
>
> *Quelques secondes suffisent pour un joueur exercé. Mais ce qui est curieux, c'est qu'il ne faut guère plus d'un dixième de seconde à un chimpanzé pour atteindre la bonne lettre.*
>
> *Bizarre... Pourquoi est-il tellement plus rapide que nous ?*
>
> *Quand nous jouons à ce jeu, nous imaginons la lettre* **S** *qui se déplace, qui tourne, qui fuit. Et quand cette image mentale mobile rattrape l'image fixe aperçue sur l'écran, nous avons gagné. Nous utilisons donc la possibilité de transporter mentalement les images, de leur faire subir certaines actions : rotations, déplacements, etc. Ces actions-là ont entre elles des relations très particulières ; les géomètres en ont fait l'inventaire ; et cet inventaire, ils l'appellent* ***groupe****.*
>
> *C'est ainsi que le groupe est antérieur, dans notre pensée, à d'autres catégories que nous pourrions croire primitives, comme « le nombre » ou « l'espace ». Le groupe spatial ? Si les singes et les hommes savent le manipuler aussi vivement, c'est qu'il s'agit d'un outil disponible à un niveau très primitif de la pensée ; peut-être est-il "câblé" quelque part dans notre cerveau, comme dans celui des animaux qui possèdent une compétence spatiale analogue à la nôtre.*

Et bien sûr, la grande idée de Poincaré (si souvent reprise en neurosciences) sur le rôle de la manipulation des objets solides et des changements de points de vue dans notre rapport à l'espace :

> *Mais il [l'esprit] commence à étudier expérimentalement les lois suivant lesquelles se composent les déplacements. L'expérience lui apprend qu'ils se comportent comme les substitutions d'un groupe d'ordre [nous dirions, de dimension] 6. [...] Mais il faut bien s'entendre. L'expérience nous apprend seulement que les déplacements se comportent à peu près comme les substitutions d'un groupe d'ordre 6. Ce n'est donc pas l'expérience qui nous fournit la notion de groupe. Cette notion préexiste ou plutôt ce qui préexiste dans l'esprit c'est la puissance de former cette notion. L'expérience n'est pour nous qu'une occasion d'exercer cette puissance. Elle nous apprend que parmi tous les groupes simples que nous pouvons former, c'est un certain groupe*[29] *qui s'écarte le moins de l'observation.*

J'espère que ce que j'ai dit de quelques-unes des structures qui nous permettent de traiter les informations sensorielles fait voir que ce n'est pas qu'une déclaration de principe ; qu'il n'est pas impossible qu'il y ait assez de résultats précis (et quantitatifs) disponibles sur les aires cérébrales pour que l'utilisation des groupes pour parler de ces structures, ou des modèles qui permettent de les décrire, puisse avoir plus qu'une visée métaphysique.

29. Poincaré discute longuement du choix en faveur du groupe des déplacements d'Euclide, et du rôle du fait que le sous-groupe des translations y soit distingué pour contribuer à la notion intuitive de point.

Daniel Bennequin contribue depuis plusieurs années à mettre les groupes au coeur de modèles biologiques précis et quantitatifs. Voici comment s'ouvre [7] :

> *Poincaré observed that the perception of space is based on active movements, and relies on the notions of invariance, covariation between sensors and environment, and active compensation. The research of Piaget has proved the importance of various kinds of geometrical invariance in cognitive and behaviorial development. To him intelligence is a form of adaptation, the continuous process of using the environment for learning.*
>
> *Adaptation is a process that can happen at the scale of evolution, development or functioning. In ecology, or in population biology and genetics, it means the adjustment or change in behavior, physiology, and structure of an organism to become more suited to an en-vironment, thus better fitted to survive and passing their genes on to the next generation (Darwin plus Mendel).*
>
> *In Neuroscience it often means the decline in the frequency of firing of a neuron in response to constantly applied environmental conditions, or more generally, any change in the relationship between stimulus and response that is induced by the level of stimulus. Adaptation is an ubiquitous essential property of sensory and motor processing, allowing the living systems to sense and anticipate what is changing in the world. As we will see, invariance can contribute to adaptation, and adaptation can create new invariance structures. Gibson gave a precise formulation of invariance and adaptation in psychology, with special emphasis on vision. From the formal point point of view, the mathematical theory of groups, and its many extensions in algebra and analysis [...], offer a clear mathematical basis for discussing the notions of invariance.*

Quelques exemples frappants concernent la planification des mouvements de nos mains et la locomotion [8, 47], où de remarquables lois reliant la vitesse instantanée et la courbure des mouvements naturels (loi de la puissance 1/3) sont abordées à l'aide des invariants différentiels du groupe équiaffine (comme chez Élie Cartan), des groupe affine et euclidien. Signalons aussi les invariances continue, projective, euclidienne qui semblent se faire jour dans fonctionnement du cortex visuel, de l'aire MT^+ dédiée au mouvement, du système parahippocampal qui abrite les célèbres cellules de place et de grille ([7], section 5).

L'idée d'utiliser les représentations du groupe de Galilée pour aborder le fonctionnement des neurones qui traitent les informations vestibulaires, formulée dans la note [9], est ce sur quoi repose la deuxième partie de cette thèse.

⋆

Dans ce qui suit, je ne peux pas, je ne veux pas et je ne dois pas aborder les questions évoquées jusqu'ici autrement que *par* les mathématiques (en les convoquant pour parler des modalités sensorielles ou des modèles qui cherchent à les décrire) ou *pour* les mathématiques (en partant des traits mathématiques de ces modèles afin de les généraliser ; c'est ce que je ferai le plus souvent).

Il y a trois ans, Daniel Bennequin me proposait de réfléchir sur le rôle du groupe de Galilée dans le traitement de l'information vestibulaire, notamment parce que le traitement de l'information *visuo*-vestibulaire est mal compris et qu'il est raisonnable de penser que la structure du groupe de Galilée inhomogène est particulièrement adaptée pour décrire les opérations effectuées par les neurones qui participent à l'intégration conjointe de ces deux modalités sensorielles − tandis que celle de la partie homogène du groupe de Galilée est probablement très adaptée pour le traitement de l'information vestibulaire "seule". Les représentations unitaires irréductibles du groupe de Galilée étaient susceptibles de fournir des ondelettes généralisées analogues aux filtres de Gabor utilisés par les neurones de V1, peut-être au moyen d'un principe d'incertitude que ces ondelettes satureraient.

Le manuscrit que vous êtes en train de parcourir, bien qu'il y reste quelques traces du projet initial, n'aborde pas cette question. Nous retrouverons cependant le système visuel dans la première partie, le système vestibulaire dans la deuxième, et naturellement les groupes, leurs représentations, l'analyse harmonique invariante seront omniprésents.

2 Représentations de groupes de Lie : quelques rappels historiques.

Le point de départ de ce travail est le fait que les constructions théoriques proposées ces dernières années pour appréhender les cartes d'orientation du cortex visuel primaire font un usage crucial d'arguments de symétrie. L'invariance des modèles par une action du groupe des déplacements du plan (pas toujours la même) est l'ingrédient principal d'une bonne partie de leur succès.

Les représentations de groupe sont l'outil idéal pour illuminer le rôle des arguments de symétrie dans un modèle, et pour l'adapter à des symétries différentes ; elles joueront le premier rôle dans cette thèse. J'expliquerai au chapitre 3 comment chaque représentation unitaire irréductible de dimension infinie du groupe des déplacements du plan abrite un objet aléatoire qui reproduit plutôt bien la géométrie des cartes d'orientation du cortex visuel primaire, et le reste de ce chapitre montrera comment on peut utiliser les représentations unitaires d'autres groupes pour obtenir des structures en pinwheels sur des espaces courbes. Dans la deuxième partie, je tenterai d'utiliser les représentations unitaires irréductibles du groupe de Galilée pour aborder la question **(a)** de la section 1.1 lorsqu'elle concerne certains neurones traitant l'information vestibulaire. Le sujet de la troisième partie de ma thèse est, quant à lui, un problème *interne* à la théorie des représentations de groupes de Lie réductifs réels.

Il est donc utile que je dise maintenant un mot des représentations de groupes de Lie, que j'explique pourquoi il est naturel que leur théorie joue ce rôle de réservoir de modèles. Elle est parfois réputée abstraite ; mais compte tenu de son histoire et de sa structure, affirmer qu'elle peut nous aider à appréhender ce que j'ai dit du cerveau est loin d'être fantaisiste.

Je vais donc rappeler quelques-uns des épisodes du développement de la théorie des représentations, et cela me permettra de présenter le problème auquel est consacrée la troisième partie de ma thèse. Le rôle de cette partie de mon introduction est ainsi

– de rappeler que la théorie des représentations fournit *par nature* des outils pour des questions très pratiques, et de ne pas laisser oublier qu'elle s'est largement construite autour de ce rôle. Dans les épisodes que je vais retracer, c'est lorsqu'elle devient *prédictive*, lorsqu'elle fournit des grandeurs et constructions *explicites* (et utilisables dans les applications, à la physique en premier lieu) à partir de la seule structure du groupe, que la théorie des représentations joue pleinement son rôle.

– De rappeler d'où vient la notion de contraction qui est essentielle pour la dernière partie de ma thèse, celle qui est purement mathématique. Pour cela, le détour par la physique est important. Nous allons d'ailleurs voir que la plupart des acteurs majeurs du développement de la théorie des représentations de dimension infinie des groupes de Lie sont partis de problèmes de physique.

Ce qui suit *n'est pas* une introduction à l'histoire du sujet, qui n'aurait pas sa place ici et qu'une quinzaine de pages serait très insuffisante à retracer : Mackey [42] ou Hawkins [23], entre autres, le font longuement et admirablement ; l'ouvrage collectif [4], en plus de bien présenter beaucoup d'outils de la théorie des représentations, contient de nombreuses remarques historiques. Mais pour comprendre le rôle de la physique dans le développement de la théorie, nous allons retrouver certains des débuts de la physique quantique. Il y aura plusieurs remarques plus biographiques que mathématiques dans ce qui suit : je crois qu'elles ne sont pas qu'amusantes, et elles m'ont été utiles pour comprendre comment le sujet s'est développé ; je choisis donc de ne pas refuser de les inclure. Mon récit est presque chronologique dans les paragraphes 2.1 à 2.6 ; j'isole deux thèmes qui sont importants dans ma thèse pour en raconter quelques développements aux paragraphes 2.7 et 2.8.

Je sais que ces discussions (ou anecdotes) historiques ne sont pas du goût de tous et que certains de mes lecteurs les jugeront légères ; pour qu'il soit possible d'aller plus vite voir le contenu de ma thèse, j'ai essayé de proposer un raccourci en détachant par des encadrés bleus les idées et les résultats, tous très célèbres, qui sont essentiels pour la suite.

2.1 1896, à Berlin : les représentations apparaissent pour comprendre un déterminant

L'étude des représentations de groupes commence avec une question de Dedekind à Frobenius. Soit G un groupe fini ; introduisons une indéterminée X_g pour chaque élément de G, notons $G = \{g_1, ...g_n\}$ et formons la matrice $\left(X_{g_i g_j}\right)_{1 \leq i,j \leq n}$, à coefficients dans $\mathbb{C}[X_{g_1}, ...X_{g_n}]$. Observons son déterminant ; c'est un polynôme en n indéterminées, il est homogène, de degré total n. Dedekind joue à décomposer ce polynôme en produit de facteurs irréductibles, pour des groupes G particuliers. Des régularités le surprennent : il y a toujours autant de facteurs irréductibles que de classes de conjugaison dans G, et chaque facteur irréductible apparaît avec une multiplicité égale à son degré. Il écrit à Frobenius en 1896 pour lui dire qu'il ne sait pas le montrer, et qu'il ne comprend pas la signification des degrés-multiplicités qui apparaissent ici. La suite de l'histoire est joliment racontée dans [20], paragraphe 4.11 : Frobenius, lassé des longs calculs sur les fonctions théta auxquels il s'adonnait jusqu'alors et avide d'un nouveau sujet de recherche, se met au travail. Il lui faut moins d'un an pour jeter les fondements de la théorie, pour reconnaître que les exposants de Dedekind sont les dimensions des espaces qui portent des représentations irréductibles, et pour prouver que Dedekind avait vu juste.

2.2 1924-30, à Berlin, à Göttingen et à Zurich : Hermann Weyl.

Frobenius tenait beaucoup à ce que son travail soit à l'écart des applications, et avait par ailleurs une animosité signalée pour les mathématiques de Lie et celles de Klein. Pendant vingt ans, la théorie des représentations sert (surtout entre les mains de Burnside) à élucider la structure des groupes finis. Schur était son élève et se tint à l'écart des groupes continus dans un premier temps, mais la géométrie algébrique et la théorie des invariants finirent par l'amener (pour comprendre les travaux d'Hurwitz et de Molien) à l'étude des représentations du groupe des rotations en 1924. Grâce à l'intégration invariante qui deviendrait célèbre une fois que Haar en aurait dégagé le cadre général, les preuves pouvaient s'adapter facilement, et Schur put montrer, par exemple, que chaque représentation du groupe des rotations est somme directe de représentations irréductibles.

Ses résultats allaient rapidement être généralisés.

En 1924, Hermann Weyl était déjà célèbre ; ses débats avec Einstein puis Study sur l'unification de l'électromagnétisme et de la gravitation l'avaient poussé à étudier le calcul tensoriel en détail et à chercher à comprendre ses fondements théoriques. Il en était à publier *Les fondements du calcul tensoriel par la théorie des groupes*, où le rôle principal est joué par les représentations irréductibles de dimension finie de $SL_n(\mathbb{C})$ et par celles du groupes symétrique. Son rapport d'activité pour l'année universitaire 1923-1924 contient les phrases suivantes :

> *The theory of finite groups, to which I was originally drawn by the theory of relativity, is more and more becoming my real area of work ; as a result of the abundance of questions that arise there, I am again strongly inclined to work on individual, purely mathematical problems. Recently, I have been considering the invariant theories associated to the most important linear groups.*

Weyl connaissait les résultats de Schur sur les groupes finis, savait l'importance des relations d'orthogonalité des caractères de représentations irréductibles pour son travail ; ils s'écrivirent beaucoup en 1924. Weyl trouva vite la formule aujourd'hui célèbre (voir plus loin, chapitre 4, paragraphe 3.4) pour les caractères des représentations irréductibles de groupes de Lie compacts. Le grand article paru en 1925 (voir [19]) qui s'occupe de cette question part du fait qu'on peut donner facilement une forme intégrale à l'orthogonalité des caractères : si χ_1 et χ_2 sont les caractères de deux représentations irréductibles non équivalentes d'un groupe compact G (ce sont donc des fonctions continues à valeurs complexes sur G),

$$\int_G \chi_1 \bar{\chi}_2 = 0$$

tandis que l'intégrale[30] de $|\chi_1|^2$ vaut 1 . Or, Weyl avait grandi près d'Hilbert ; il l'avait vu partir des relations d'orthogonalité entre fonctions trigonométriques pour dégager ce que nous appelons aujourd'hui *base hilbertienne*, et construire la théorie spectrale. Ce fut un pas de géant lorsqu'il comprit qu'écrire la formule ci-dessus, si G est le cercle[31], c'était redécouvrir l'analyse de Fourier.

La théorie des représentations du cercle, c'est celle des séries de Fourier.

Bientôt Weyl serait amené par la physique à s'intéresser au cas où G est la droite réelle (abélienne, mais pas compacte), et renforcerait ce slogan. Ses grands théorèmes sur les groupes compacts (et non abéliens) doivent beaucoup à ce point de vue : dans le travail avec Fritz Peter[32] qui sera publié en 1927, l'idée que la théorie des représentations de groupes *généralise* l'analyse de Fourier, et que cela doit être un guide pour trouver les théorèmes, joue le rôle principal. À la page 5 de l'article de Peter et Weyl, elle sert explicitement de point de départ. Puisqu'il jouera un rôle pratique au chapitre 6, et bien qu'il s'agisse presque du seul résultat très élémentaire de théorie des représentations dont je vais donner un énoncé complet, je rappelle le théorème principal de Peter et Weyl.

Soit G un groupe compact. Si $T : G \mapsto GL(V)$ est une représentation (unitaire) irréductible, choisissons une base $\{e_i\}$ de V (elle est finie), et formons la collection $\mathcal{C}_T = \{\langle Te_i, e_j\rangle\}_{i,j=1\ldots\dim(V)} \in \mathcal{C}(G)^{\dim(V)^2}$ de fonctions sur G qui donne les coefficients des matrices de $T(G)$ dans la base $\{e_i\}$. Alors

• Le sous-espace vectoriel de $\mathcal{C}(G)$ engendré par $\mathcal{C}_T$ ne dépend que de la classe d'équivalence (que je noterai $[T]$) de T. Ce sous-espace ne dépend donc pas du choix de la base $\{e_i\}$.

• Notons $\hat{G}$ l'ensemble des classes d'équivalence de représentations irréductibles de G. La famille $\bigcup_{[T]\in\hat{G}} \mathcal{C}_T$ est une base hilbertienne de $\mathbb{L}^2(G)$.

---

30. Les résultats de Haar ne viendraient qu'en 1933, mais pour un groupe de Lie, on peut utiliser une forme différentielle pour intégrer, c'est ce que fait Weyl.

31. Dans ce cas χ est de la forme $g = e^{ix} \mapsto e^{inx}$, avec n entier.

32. On sait très peu de choses sur Fritz Peter, qui semble avoir soutenu sa thèse en 1923 et quitté la recherche pour être directeur d'une école secondaire immédiatement après avoir travaillé avec Weyl.

Nous venons de voir la parenté entre l'approche de Peter et Weyl et l'étude des propriétés spectrales des opérateurs auto-adjoints. Il se trouve que ces dernières sont plus qu'en vogue en 1927 : la mécanique quantique naissante en fait la clé de nos observations. Une quantité physique "observable" devient un opérateur auto-adjoint sur un espace de Hilbert $\mathcal{H}$, ses vecteurs propres les états du système pour lesquels la valeur de l'observable physique est bien déterminée, et les valeurs propres correspondantes les résultats numériques des mesures sur ces états. Weyl se tient au courant, bien sûr : comment ne le ferait-il pas, lui qui venait de la physique, et que Schrödinger, collègue de Zürich et ami très proche, remercie dans son premier article sur la mécanique ondulatoire [55] ?

Des opérateurs qui jouent un rôle privilégié et dont le spectre détermine l'objet qui est au centre des préoccupations, il y en a aussi dans sa théorie du plus haut poids pour les groupes de Lie semi-simples compacts. Il regarde : *les opérateurs sont les mêmes*[33]. Les mêmes ! Aussitôt Weyl publie [67] ; en voici les premières phrases.

> *En mécanique quantique, on peut distinguer clairement deux questions :*
>
> 1. *Comment puis-je trouver la matrice hermitienne qui représente une grandeur donnée pour un système physique dont la constitution est connue ?*
> 2. *Une fois que je connais cette matrice, quelle est sa signification physique, quel type d'affirmations physiques puis-je en tirer ?*
>
> *Von Neumann a répondu récemment à la seconde question d'une manière claire et profonde...*

... mais sur la première question, Weyl "*croit avoir fait quelques progrès au moyen de la théorie des groupes*". Son outil est le suivant : la représentation de l'algèbre enveloppante de l'algèbre de Lie qu'on peut associer à une représentation unitaire ou projective d'un groupe de Lie fournit des opérateurs auto-adjoints, et *ce sont ces opérateurs qui sont d'excellents candidats à représenter les quantités physiques.*

L'exemple choisi par Weyl dans [67] est celui des opérateurs *position* et *impulsion*, et puisque le spectre de l'opérateur position doit être l'espace tout entier, cela le mène à des espaces de dimension infinie. Retenons deux choses que fait Weyl dans [67] :

(i) Signaler qu'il est avantageux, lorsqu'on a affaire à un opérateur auto-adjoint non borné A sur un espace de Hilbert, d'étudier la représentation unitaire de $\mathbb{R}$ donnée par $t \mapsto \exp(itA)$, et comprendre le lien entre le théorème spectral pour A, les représentations unitaires de $\mathbb{R}$ et l'analyse de Fourier.

(ii) Étudier les représentation *projectives*[34] de $\mathbb{R}^6$ sur $\mathbb{L}^2(\mathbb{R}^3)$.

Le point (ii) deviendra vite le théorème de Stone-von Neumann (voir la section 2.5 ci-dessous), et le point (i) est à ma connaissance le premier traitement des représentations unitaires du groupe des translations de $\mathbb{R}$.

Nous sommes toujours en 1927 ! Ce travail de Weyl, et les premières études de Wigner que je vais évoquer bientôt, sont un nouvel éclair dans l'étrange tempête que traverse la physique. Pour beaucoup de physiciens, les matrices et la réduction des endomorphismes

33. Exemple, pour le groupe des rotations : l'opérateur de Casimir, puisqu'il est dans le centre de l'algèbre enveloppante, agit comme une homothétie dans l'espace de chaque représentation irréductible, et la valeur propre indique le plus haut poids de la représentation. Cet opérateur n'est autre que celui qui donne le "moment cinétique total" ; voir [66], II.5.

34. Voir le début du chapitre 5.

auto-adjoints, c'était déjà beaucoup ; certains (le premier semble avoir été Ehrenfest) commencent à utiliser le vilain terme de “Gruppenpest" pour parler de l'influence imprévue de la théorie des groupes en mécanique quantique. Il faudra dix ans pour que le rôle des groupes soit clair dans beaucoup d'esprits.

Mais laissons un instant la physique ; il va nous falloir deux détours par Moscou pour comprendre comment la théorie des représentations de groupes non-abéliens est devenue en deux temps un sujet mathématique de premier plan.

2.3 1930-1939, à Moscou, à Princeton et à Paris : la théorie générale des caractères des groupes abéliens

Pontryagin, encore étudiant, était déjà topologue. Il avait dix-neuf ans en 1927, et autour d'Alexandrov qui le guidait dans ses études, il travaillait sur les relations de dualité entre l'homologie (à coefficients entiers) d'un sous-espace de l'espace euclidien et la cohomologie de son complémentaire. Autour de 1933, il chercha à changer les coefficients pour les groupes d'homologie. Cela lui révéla une propriété des groupes abéliens et de leurs caractères, tout à fait indépendente de l'origine topologique du problème : associer l'ensemble de ses caractères à un groupe abélien donné met en dualité les groupes abéliens *discrets* dénombrables et les groupes abéliens *compacts*.

Pour voir d'où vint la suite, il nous faut une question venue de l'analyse : Bochner avait étudié les fonctions presque-périodiques avec Harald Bohr dès leur définition en 1923, et contribué au passage à clarifier bien des notions de l'analyse de Fourier sur la droite réelle. Salomon Bochner dut s'exiler de Münich à Princeton en 1933, et une collaboration avec von Neumann commença ; elle visait à faire entrer les fonctions presque-périodiques de Bohr dans le cadre de la théorie des groupes et de celles des représentations. Von Neumann, qui était bien sûr au courant des travaux de Weyl, vit tout de suite l'intérêt des nouveautés venues de Moscou [35]. Van Kampen, récemment arrivé aux Etats-Unis, travaillait aussi sur les fonctions presque-périodiques : sur la suggestion de von Neumann, il étendit la dualité de Pontryagin à tous les groupes *abéliens localement compacts*.

Autour de 1935, André Weil introduisit ce que nous appelons aujourd'hui la compactification de Bohr pour faire le point sur ces résultats, et les fusionner avec le point de vue de Weyl sur les séries de Fourier. Son exposé dans *L'intégration dans les groupes topologiques et ses applications à l'analyse* parut en 1940 ; c'est, je crois, le premier exposé systématique d'éléments de la théorie des représentations de groupes non compacts, et c'est encore l'une des meilleures introductions au sujet.

Nous voilà bien loin de la physique. Mais l'étude des fonctions presque-périodiques a cimenté le lien avec l'analyse de Fourier classique (y compris lorsqu'elle est fine avec Bochner) et mené à la théorie *générale* des représentations de groupes abéliens, sous la forme exposée par Weil. Voilà comment Weil, Pontryagin, Bochner, von Neumann et van Kampen avaient donné une généralité et une profondeur nouvelles à la remarque de Weyl :

La théorie des représentations de groupes abéliens, c'est l'analyse de Fourier.

35. Pontryagin avait lui-même signalé, dans une note de 1934 aux Comptes-Rendus [50], le lien entre la théorie des fonctions presque-périodiques et celle des caractères des groupes abéliens. L'article de von Neumann de 1934 [63] ne le cite pas, mais dit clairement ce qu'il doit aux résultats de Peter et Weyl ; par conséquent, je ne sais pas si ce que je viens de dire est historiquement exact, ou si Von Neumann a découvert le lien indépendamment.

2.4 1939 et 1945, à Princeton : Wigner, Dirac et les particules élémentaires

En 1926, Wigner travaillait dans la tannerie de son père, et il avait appris le génie chimique à Berlin : son père avait insisté, ce serait plus formateur que des études de physique pour celui qu'il destinait à reprendre l'entreprise familiale. Mais il avait passé son temps libre à Berlin à écouter des exposés de physique et de mathématiques. Diplôme en main, il avait tenu parole à contrecoeur et était revenu à Budapest un an auparavant ; il pensait que c'était pour de bon. Une lettre, soudain, vient de Berlin : elle est signée Weissenberg [36] de l'institut de cristallographie ; Wigner peut-il venir l'aider à apprendre la théorie des groupes pour comprendre ses applications à la cristallographie ?

Son père accepte, Wigner revient. Il revoit un autre jeune hongrois, un ami d'enfance et de lycée, qui avait été apprendre le génie chimique à Berlin à la demande de ses parents lui aussi, et qui lui aussi suit avec attention les débuts de la physique quantique : c'est von Neumann. La suggestion ne tarde pas à venir : il y a des matrices chez Heisenberg et Wigner a envie d'utiliser la théorie des groupes, pourquoi ne pas regarder du côté des représentations ?

Wigner y trouvera des trésors. Son premier article paraît en 1927, à peu près en même temps que celui de Weyl ; il s'occupe de problèmes bien plus spécifiques que ce dernier et je ne les évoquerai pas ; mais, moins éthérées que celles de Weyl, ses idées contribuent grandement à répandre la "peste des groupes" chez les physiciens.

Vient 1930, le départ à Princeton en compagnie de von Neumann [37]. Wigner est physicien, travaille sur les noyaux atomiques et sur ce que la mécanique quantique a à dire de la physique des solides, mais il y retrouve les symétries qui feront sa gloire. Partout les groupes et leurs représentations l'aident.

Ses idées sur les équations imposées aux fonctions d'onde et sur leurs symétries mûrissent. Pendant ce temps, de mystérieux articles de Majorana (1932) et Dirac (1936) proposent des équations d'onde qui sortent presque de nulle part ; Dirac dit [17] :

> *The elementary particles known to present-day physics, the electron, positron, neutron, and proton, each have a spin of a half, and thus the work of the present paper will have no immediate physical application. All the same, it is desirable to have the equations ready for a possible future discovery of an elementary particle of spin greater than a half.*

Le mot "groupe" n'apparaît dans aucun de ces deux articles, mais l'ingrédient crucial est de requérir que les équations soient linéaires et invariantes par les translations et les transformations de Lorentz, puis − comme Dirac l'avait fait avec les spineurs pour l'électron − d'ajouter des dimensions au but.

Wigner comprend qu'ils sont en train de découvrir des espaces qui portent des représentations irréductibles du revêtement universel du groupe de Poincaré − celui des transformations *affines* de $\mathbb{R}^4$ qui préservent la métrique de Minkowski et assemble ainsi les *translations* de $\mathbb{R}^4$ et les transformations *linéaires* de Lorentz (celles de $SO(3,1)$ et de son revêtement universel $SL_2(\mathbb{C})$). Il saute à pieds joints par-dessus ce qui le sépare des mathématiciens : il va partir de la structure abstraite groupe et les trouver toutes !

Son étude [68] de 1937 − parue en 1939 [38] − est, à l'exception peut-être des travaux

36. Ce sont les amis de Wigner qui lui ont suggéré la démarche.

37. L'invitation à Princeton semble avoir eu pour objectif caché d'aider von Neumann à franchir l'Atlantique en lui gardant un ami... !

38. L'article de Wigner a été rejeté comme "sans intérêt pour les mathématiques" par une première revue ; von Neumann a dû intervenir pour lui en trouver une autre.

autour du théorème de Stone-von Neumann [39], la première qui s'attaque aux représentations irréductibles *de dimension infinie* d'un groupe non abélien.

Laissons Sternberg nous raconter les conséquences de ce travail pour la physique [59] :

All of the recent theories of elementary particles have been shaped by the paper by Wigner, containing the classification of the irreducible representations of $SL(2,\mathbb{C}) \ltimes \mathbb{R}^{1,3}$ described above. It is difficult to overestimate the importance of this paper, which will certainly stand as one of the great intellectual achievements of our century. It has not only provided a framework for the physical search for elementary particles, but has also had a profound influence on the development of modern mathematics, in particular the theory of group representations. From our point of view, we can summarize Wigner's main points as follows. The logic of physics is quantum mechanics. Hence, a symmetry group of the system manifests itself as a unitary (or possibly anti-unitary) representation. Ignoring the anti-unitaries (for example, by considering connected groups), one posits that

(i) *an elementary particle "is" an irreducible unitary representation of the group G of physics, where these representations are required to satisfy certain physically reasonable restrictions, and where*

(ii) *the group G of physics is [...] the universal covering of the Poincaré group.*

[...] In the past 50 years or so since Wigner's work various modifications have been introduced. Even at the time of publication of Wigner's article, the scheme was not completely satisfactory. In the intervening years various other invariant "quantum numbers" of particles have been registered, such as isotopic spin, strangeness, etc. The main thrust of recent theories (or at least most of them) has been to modify (ii) by enlarging the group G. The most successful of such theories, involving quarks and electroweak unification, indicates that the group $SU(3)$ should be contained, in some way, in the group G. Recent speculation concerning the so-called supersymmetries suggests the notion of a group might have to be slightly enlarged. However, point (i) in the dogma remains unchanged.

Pour prédire les équations d'onde linéaires qui servent à la physique,
on peut partir des symétries qu'on leur souhaite,
et observer les représentations linéaires du groupe correspondant.

Mais nous sommes en 1939 − Wigner est l'un des initiateurs du projet Manhattan (c'est lui qui suggéra de prévenir Roosevelt du danger, et qui accompagna Szilard chez Einstein). L'étude systématique des représentations de dimension infinie attendra la fin de la guerre.

2.5 1947, entre Moscou, Princeton et Harvard : début de la théorie générale

L'étude des représentations irréductibles de groupes non-compacts (et non-abéliens) commence brusquement sitôt la guerre achevée. Quatre articles parus en moins de deux ans ont des titres étrangement similaires :

– (1945) *Unitary representations of the Lorentz group*, par Paul Dirac de Cambridge, remarque que la théorie quantique de l'oscillateur harmonique [40] suggère l'existence de représentations unitaires de dimension infinie du groupe de Lorentz $SO(3,1)$. Dirac les appelle *expanseurs* : ce sont "une nouvelle sorte de tenseurs, avec un nombre infini de composantes et dont la longueur peut être exprimée par une expression définie positive". Dirac ne cite pas son beau-frère Wigner, et leurs résultats semblent sans lien.

39. Interprétation anachronique : on ne comprendra qu'après la guerre qu'ils parlent des représentations projectives de $\mathbb{R}^2$, voir plus loin.

40. Celle de Fock.

- (1947) *Unitary representations of the Lorentz group*, par Israel Mossievitch Gelfand et Mark Aronovitch Naimark de Moscou. C'est Gelfand qui a, semble-t-il, voulu le premier comprendre les représentations de groupes non-compacts de façon systématique. Cet article suit la suggestion de Dirac de passer au revêtement universel (c'est Dirac qui avait vu qu'obtenir des spins demi-entiers nécessite ce passage pour le groupe des rotations), et s'occupe des représentations unitaires irréductibles de $SL_2(\mathbb{C})$ de façon systématique ; il introduit la *série principale*. Je présenterai ce travail au prochain paragraphe.
- (1947) *Infinite irreducible representations of the Lorentz group*, par Harish-Chandra de Princeton, est la thèse d'un étudiant de Dirac venu de l'Inde. Lorsqu'on parcourt le court article de Dirac, aucun des développements mathématiques futurs n'est reconnaissable. Comme nous allons le voir, les précisions données par Harish-Chandra sont les premiers pas sur un chemin qui allait le mener à ce que sa personne se confonde avec la théorie des représentations de groupes de Lie semi-simples.
- (1947) *Irreducible unitary representations of the Lorentz group*, par Valentine Bargmann de Princeton. Bargmann était l'assistant d'Einstein, et travaillait régulièrement avec Einstein, Wigner (ils ont écrit ensemble une sorte de suite à l'article de 1939, où figurent les équations d'onde explicites, en 1948), et von Neumann [41]. Bargmann signale qu'il en a obtenu les résultats entre 1940 et 1942, et en plus de Wigner et von Neumann, il y remercie Pauli pour lui avoir suggéré d'aborder ce problème. Parmi les quatre articles, c'est le seul qui fasse explicitement référence au travail de Wigner. C'est aussi le plus complet et, de très loin, le plus abondamment cité aujourd'hui : nous allons voir pourquoi [42].

L'article de Dirac est court comme toujours, et pour ce qui est mathématique, il se contente de signaler l'existence d'équations aux dérivées partielles dont l'espace des solutions porte une représentation de dimension infinie du groupe de Lorentz. Son rôle dans le développement de la théorie semble avoir été surtout de susciter la curiosité des auteurs des articles parus en 1947. Pour Harish-Chandra la filiation est directe ; quant à Gelfand et Naimark et à Bargmann, tous prennent le temps de montrer que les représentations identifiées par Dirac sont en fait réductibles et d'identifier leurs facteurs irréductibles.

Les trois autres articles, en revanche, font brusquement du sujet un thème d'avenir pour les mathématiques. Nous reviendrons en 2.6 sur le travail d'Harish-Chandra et ses suites ; observons d'abord les autres avancées de 1947.

⋆

Le long texte de Gel'fand et Naimark se fixe pour objectif explicite d'étudier *toutes* les représentations unitaires de $SL_2(\mathbb{C})$, en déterminant *toutes* les représentations unitaires irréductibles à équivalence unitaire près ; ses auteurs sont au courant de l'importance de la question pour la physique et se réfèrent bien à l'article de Dirac, mais c'est en mathématiciens qu'ils abordent la question. Leur travail, très complet (90 pages), sera d'ailleurs suivi de près d'une monographie sur les représentations de groupes classiques (1950), et de très nombreux résultats généraux sur les représentations unitaires sont en train d'être dégagés par l'école de Gelfand, si bien qu'il est tentant d'attribuer à l'école russe le projet d'une théorie générale pour les groupes non-compacts. L'introduction contient la phrase : "les méthodes utilisées ici peuvent être appliquées à tous les groupes semi-simples complexes".

41. Notamment sur le traitement algorithmique de probèmes sur les matrices de grande dimension, ce qui ne peut qu'impressioner aujourd'hui.

42. Le talent de Bargmann pour écrire n'y est peut-être pas pour rien : lue en 2015, l'introduction de son long article n'a pas pris une ride.

En reprenant les notations de Gelfand et Naimark, soit D le sous-groupe (abélien) de $G = SL_2(\mathbb{C})$ formé des matrices diagonales. À chaque *caractère unitaire* de D (chaque morphisme de D vers le groupe des nombres complexes de module 1), Gelfand et Naimark associent une représentation unitaire irréductible de G. Un tel caractère χ est de la forme $\begin{pmatrix} re^{i\theta} & 0 \\ 0 & r^{-1}e^{-i\vartheta} \end{pmatrix} \mapsto e^{ik\vartheta}r^{i\rho}$, avec $(k, \rho) \in \mathbb{Z} \times \mathbb{R}$. Lorsque f est un élément de $\mathbf{L}^2(\mathbb{C})$ et $g = \begin{pmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{pmatrix}$ est un élément de $SL_2(\mathbb{C})$, posons

$$U_{\chi,g}f = z \mapsto |-g_{12}z + g_{22}|^{-2-i\rho} \left(\frac{-g_{12}z + g_{22}}{|-g_{12}z + g_{22}|} \right)^{-n} f\left(\frac{g_{11}z - g_{21}}{-g_{12}z + g_{22}} \right). \tag{2.1}$$

Alors $U_{\chi,g}$ définit un opérateur unitaire sur $\mathbf{L}^2(\mathbb{C})$, et $g \mapsto U_{\chi,g}$ est une représentation unitaire de G ; Gelfand et Naimark démontrent qu'elle est irréductible. Ils appellent *série principale* la famille de représentations ainsi obtenue. La suite de leur long texte suit un programme aujourd'hui familier : ils étudient les équivalences entre ces représentations, leurs caractères ; puis ils démontrent qu'il est possible de décomposer la représentation régulière de G à l'aide des représentations de série principale, écrivent la formule de Plancherel correspondante. Ils s'attaquent ensuite à l'ensemble des représentations unitaires irréductibles, et montrent qu'en autorisant ρ à avoir une partie imaginaire, il est possible d'étendre un peu la série principale au cas où $n = 0$ et $-2 < \mathrm{Im}\rho < 0$, pourvu qu'on accepte de changer le produit scalaire pour que (2.1) continue à définir des représentations unitaires ; la *série complémentaire* ainsi obtenue, lorsqu'on l'ajoute à la série principale, termine la liste des représentations unitaires irréductibles à équivalence près.

Soit K le sous-groupe de $SL_2(\mathbb{C})$ formé des matrices *triangulaires supérieures* (attention, de nos jours, *personne* ne le note ainsi) ; il s'écrit $K = DZ$, où Z est le groupe (ici abélien) des matrices triangulaires supérieures dont les coefficients diagonaux valent 1. Gelfand et Naimark insistent dès les premières sections de leur article sur l'importance du sous-groupe K, de la décomposition $K = DZ$ et du quotient G/K pour leurs résultats et démonstrations, bien que ce fait ne soit pas visible sur ce que je viens de transcrire. Nous verrons l'importance de cette remarque pour le travail d'Harish-Chandra dans le cas réel.

⋆

Bargmann étudie quant à lui les représentations du véritable "groupe de Lorentz", $SO(3, 1)$, et celles de son analogue pour un "espace-temps de dimension trois" − le groupe $SO(2, 1)$, qui est isomorphe à $SL_2(\mathbb{R})$. Les résultats pour lesquels son travail est célèbre sont ceux qui concernent $SL_2(\mathbb{R})$. Bargmann indique trois motivations pour son étude : l'intérêt possible pour la physique signalé par Dirac, le fait qu'il s'agit d'étudier les représentations de deux groupes qui ne sont pas compacts, et le fait que la connaissance des représentations de ces deux groupes est nécessaire pour classifier les représentations du groupe de Poincaré d'après Wigner.

Pour $SL_2(\mathbb{R})$, il y a une série analogue à la série principale de Gelfand et Naimark. Bargmann la découvre indépendamment, et la décrit en détail. Mais il y a d'autres représentations unitaires irréductibles de $SL_2(\mathbb{R})$: soit $\mathbb{H}$ le demi-plan supérieur dans $\mathbb{C}$ et k un entier supérieur ou égal à 2 ; Bargmann découvre que l'espace [43] $\mathcal{H}_k$ des fonctions

43. L'espace défini par Bargmann en 1947 rassemble plutôt des fonctions holomorphes sur le disque unité dans $\mathbb{C}$, c'est celui qu'on obtient lorsqu'on transfère le demi-plan sur le disque et $SL_2(\mathbb{R})$ sur $SU(1, 1)$ de la façon habituelle, celle que je rappellerai dans la section 2.3 du chapitre 3.

holomorphes f sur $\mathbb{H}$ pour lesquelles la quantité $\int_{\mathbb{H}} |f(x+iy)|^2 y^{k-2} dxdy$ est finie, muni de l'action de $SL_2(\mathbb{R})$ donnée par la formule

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot f = z \mapsto (-bz+d)^k f\left(\frac{az-c}{-z+d}\right), \tag{2.2}$$

porte une représentation unitaire irréductible de $SL_2(\mathbb{R})$. Il y a aussi une infinité dénombrable de représentations qui agissent sur des espaces de fonctions anti-holomorphes. Bargmann découvre que les éléments matriciels de ces représentations sont *de carré intégrable* (c'est le théorème 6, section 12, dans [5]), alors (théorème 7) que ce n'est pas le cas pour les autres représentations. Il appelle cette nouvelle série la *série discrète*, et montre une version adaptée de l'orthogonalité des éléments matriciels (théorème 8), puis du théorème de Peter-Weyl (théorème 9) qui lui dit que série principale et série discrète suffisent à décomposer $\mathbf{L}^2(SL_2(\mathbb{R}))$.

Une remarque inattendue prendra une importance considérable par la suite : être une fonction holomorphe qui est invariante par la restriction à $SL_2(\mathbb{Z})$ de l'action (2.2), c'est être une forme modulaire de degré k au sens classique.

⋆ ⋆ ⋆

La même année paraît une étude importante signée d'un ancien étudiant de Stone, grandi à l'école de Bochner et von Neumann. Il s'appelle George Mackey ; sa thèse de 1942 portait sur les topologies faibles associées aux dualités entre espaces vectoriels, mais avant de commencer à y travailler, il avait été étudiant en physique, et son intérêt pour la mécanique quantique ne le quitterait jamais. Autour de 1945, à Harvard, il cherche à préciser les liens entre l'analyse harmonique et les relations de Heisenberg ; il se rend compte que ce que nous appelons aujourd'hui le théorème de Stone-von Neumann n'est rien d'autre qu'une classification des représentations projectives du groupe des translations de $\mathbb{R}^2$.

Mackey fait le point sur ce sujet dans [39] ; c'est de cet article que nous vient l'attribution "Stone-von Neumann". Il y fait grand usage des décompositions en *intégrale directe* d'espaces de Hilbert que von Neumann venait de dégager, et le remercie dans son texte.

Cette étude de 1947 s'occupe de groupes localement compacts *abéliens*, comme Bochner et von Neumann. Bientôt Mackey remarque une parenté entre ses réflexions sur le théorème de Stone-von Neumann et des constructions qui apparaissent dans l'étude maintenant classique des représentations linéaires des groupes finis *non abéliens*.

La notion très générale de *système d'imprimitivité* pour une représentation de groupe localement compact (non abélien) qu'il dégage en fusionnant ce thème ancien avec ses découvertes récentes, et sa version des *représentations induites* pour les groupes localement compacts (qui permet, lorsque H est un sous-groupe fermé de G, de construire une représentation de G à partir d'une représentation de H), allait immédiatement devenir un outil indispensable pour prolonger les travaux de Gelfand-Naimark et Bargmann.

Lorsqu'il énonce ses premières idées sur le sujet dans une note de 1949 [40], il en indique une application "immédiate" mais triomphale : il peut déterminer les représentations unitaires irréductibles de produits semi-directs $H \ltimes A$ lorsque le facteur distingué A est abélien, et démontrer que la méthode utilisée par Wigner pour les représentations du groupe de Poincaré est rigoureuse et générale. Ce travail de 1949, dont je décrirai les résultats (avec démonstrations) au chapitre 5, est plus qu'important pour cette thèse : presque tous les résultats que je vais démontrer s'y rattachent d'une manière ou d'une autre.

2.6 1950 à 1975, à Princeton, le travail d'Harish-Chandra

La théorie des représentations de groupes de Lie semi-simples non compacts, qui est le cadre dans lequel la plupart des résultats de ma thèse ont été pensés ou formulés, est pour l'essentiel l'oeuvre d'Harish-Chandra. Elle est solidement établie comme un chapitre important des mathématiques "pures" ; ses liens (et ceux de ses prolongements au cas des groupes réductifs sur d'autres corps) avec la théorie des nombres fascinent depuis presque cinquante ans.

Mais Harish-Chandra a commencé physicien, et lorsqu'on est habitué à ses grands articles ultérieurs [44], le style "pédagogique" de son étude de 1947 sur le groupe de Lorentz étonne. Harish-Chandra, bien que peu bavard, a raconté sa conversion aux mathématiques :

> *Soon after coming to Princeton I became aware that my work on the Lorentz group was based on somewhat shaky arguments. I had naively manipulated unbounded operators without paying any attention to their domains of definition. I once complained to Dirac about the fact that my proofs were not rigorous and he replied, "I am not interested in proofs but only in what nature does." This remark confirmed my growing conviction that I did not have the mysterious sixth sense which one needs in order to succeed in physics and I soon decided to move over to mathematics.*

Trente-six ans plus tard, la biographie rédigée par Howe pour l'académie des sciences américaine commence ainsi.

> *Harish-Chandra was, if not the exclusive architect, certainly the chief engineer of harmonic analysis on semisimple lie groups. This subject, with roots deep in mathematical physics and analysis, is a synthesis of Fourier analysis, special functions and invariant theory, and it has become a basic tool in analytic number theory, via the theory of automorphic forms. It essentially did not exist before World War II, but in very large part because of the labors of Harish-Chandra, it became one of the major mathematical edifices of the second half of the twentieth century.*

D'excellentes références pour ce qui suit sont les synthèses de R. Herb [26], de R. Howe [27], de R. P. Langlands [35], et le livre introductif de V. S. Varadarajan [61] .

Pendant près de vingt-cinq ans, Harish-Chandra s'est concentré sur la tâche suivante. Soit G un groupe de Lie semisimple (disons, connexe et de centre fini, puisque c'était la coutume en son temps). Fixons une mesure de Haar sur G (qui est unimodulaire), et observons la représentation régulière de G sur $\mathbf{L}^2(G)$. Peut-on en identifier les composantes irréductibles ?

Pour comprendre le travail d'Harish-Chandra, énonçons une version de la *formule de Plancherel* abstraite. Lorsque π est une représentation unitaire irréductible de G, définissons un opérateur sur l'espace où π agit par

$$\pi(f) = \int_G f(g)\pi(g)dg$$

dès que f est une fonction continue à support compact sur G. Harish-Chandra a montré en 1954 que cet opérateur admet une trace. La forme linéaire $f \mapsto \text{Trace}(\pi(f))$ s'étend alors en une *distribution* Θ_π sur G, qui est bien sûr l'analogue du caractère d'une représentation irréductible de groupe fini.

Il existe un sous-ensemble $\widehat{G}_{temp}$ du dual unitaire de G (le dual unitaire est l'ensemble des classes d'équivalence de représentations unitaires irréductibles de G), et une mesure μ

44. Langlands raconte [36] qu'au moment de choisir les lauréats de la médaille Fields en 1958, certains s'opposèrent à ce que deux mathématiciens étiquetés "bourbakistes" soient récompensés (l'opposant était pourtant une idole de Bourbaki). Le "bourbakiste" finalement choisi plutôt qu'Harish-Chandra fut... René Thom, dont les déclarations ultérieures sur Bourbaki rendent amusantes ces querelles d'étiquettes.

sur le dual unitaire de G dont le support est $\widehat{G}_{temp}$ qui, si δ_{1_G} est la distribution de Dirac à l'origine, donne lieu à l'égalité de distributions

$$\delta_{1_G} = \int_{\widehat{G}_{temp}} \Theta_\pi d\mu(\pi). \tag{2.3}$$

Il a fallu vingt-cinq ans de travail à Harish-Chandra pour décrire $\widehat{G}_{temp}$ et μ. Langlands rapporte [35] :

> *It appears that by the early 1950s he had already glimpsed the outlines of the theory of harmonic analysis on real semisimple groups, and in the next ten years he marched towards it with formidable determination and resourcefulness, inventing techniques and constructions as he advanced. Even after the wave of advance had crested in the discrete series and its force been partly diverted into other channels, the tenacity in the search for solutions to technical difficulties which was a characteristic of Harish-Chandra's style remained.*

⋆

Langlands vient de nous signaler que la *série discrète* a joué un grand rôle : *isolées* dans le dual tempéré [45], les représentations correspondantes sont celles qui ont un élément matriciel *de carré intégrable*, et généralisent les représentations identifiées par Bargmann – celles où on retrouvait les formes automorphes. Il n'y en a qu'une infinité dénombrable ; mais identifier les représentations de série discrète et calculer explicitement leurs caractères était un travail immense. Harish-Chandra a mis dix ans à le mener à bien.

Harish-Chandra avait compris très tôt que les représentations tempérées viennent en *séries* et que certains sous-groupes abéliens de G, les sous-groupes de Cartan, ont un rôle crucial à jouer. Ils ont tous la même dimension, notons-la r, et il existe un entier $s \in \{0,..r\}$ tel que chaque sous-groupe de Cartan ait une composante neutre de la forme TA, où $T \simeq (\mathbf{S}^1)^q$ et $A \simeq \mathbb{R}^{r-q}$ pour un certain entier $q \in \{0,..s\}$. À chaque classe de conjugaison de sous-groupe de Cartan est associée une classe de conjugaison de sous-groupes *paraboliques*, qui sont des sous-groupes de G d'autant plus petits que q est petit, et qui jouent pour la construction des représentations de G le rôle que jouait le sous-groupe des matrices triangulaires supérieures chez Gelfand et Naimark.

La construction est la suivante : le groupe qui joue le rôle que jouait le groupe D des matrices diagonales chez Gelfand et Naimark est le centralisateur L de A dans G. Cependant, il n'est plus abélien : il s'écrit MA, où M est un sous-groupe réductif de G qui contient T et intersecte A trivialement, mais qui n'est pas abélien ni connexe en général et qui peut être très gros (c'est G tout entier si $q = r$). À la place des caractères de D, il faut utiliser les représentations de série discrète de M et les caractères unitaires de A. À la place du groupe Z, on utilise un sous-groupe nilpotent N de G qui normalise A (déterminé à un choix fini près par les autres données), et le *sous-groupe parabolique cuspidal* $P = MAN$ est l'analogue voulu du groupe des matrices triangulaires supérieures.

Pour associer une représentation unitaire de G à ces données, on étend trivialement le caractère donné de A à AN ; avec la représentation choisie de M cela fournit une représentation unitaire de P, puis on utilise la notion d'induction unitaire de Mackey pour construire une représentation de G.

Lorsque $q = 0$, la construction donne l'analogue de la série principale de Gelfand et Naimark ; l'entier s n'est égal à r que lorsque G admet une série discrète, mais lorsque $q = s = r$, elle produit la série discrète. La construction d'une représentation de G à

45. Voir le début du chapitre 8 pour la définition de la topologie sur $\widehat{G}_{temp}$.

partir des données précédentes est en fait un cas simple de l'induction unitaire (voir le début du chapitre 7 de [34]) ; quelques détails sont donnés au paragraphe 4.1 du chapitre 7.

Une fois la série discrète décrite et l'induction parabolique bien comprise, il faut encore trouver la mesure de Plancherel et prouver l'égalité (2.3). Cela nécessite une analyse très fine du comportement asymptotique des éléments matriciels des représentations tempérées de G. Je ne la décrirai pas, c'est un gigantesque travail ; vingt-quatre ans de dur labeur séparent la réalisation de ce programme pour $SL_2(\mathbb{R})$ et son analogue pour un groupe semisimple quelconque.

L'esprit de cette partie de mon introduction m'invite à inclure un mot sur la personnalité d'Harish-Chandra. Voici comment Herb conclut son texte de présentation :

Perhaps because the details of this work were so technical, Harish-Chandra liked to have a simple philosophy to guide his efforts. He says in the introduction to a survey paper, "Our entire approach to harmonic analysis on reductive groups is based on the philosophy of cusp forms." The term cusp form comes from the theory of automorphic forms and was used by Harish-Chandra as a name for the matrix coefficients of the discrete series representations which play a pivotal role in the harmonic analysis. The strong analogy between his work on real groups and the theory of automorphic forms from number theory led him to believe that whatever is true for a real reductive group is true for a p-adic reductive group. He called this the Lefschetz Principle because of its resemblance to the Lefschetz principle in algebraic geometry. It led him to a proof of the Plancherel formula in the p-adic context as well, although in this case he was unable to explicitly parameterize the discrete series. This unified theory for real and p-adic groups reinforced his belief that harmonic analysis on a semisimple group was a special thing. At the end of his 1972 lecture series at the Williamstown conference, he told a story which he attributed to Chevalley. The story relates to the time before Genesis when God and his faithful servant, the Devil, were preparing to create the universe. God gave the Devil pretty much a free hand in building things, but told him to keep off certain objects to which He Himself would attend. Chevalley's story was that semisimple groups were among the special items. Harish-Chandra added that he hoped that the Lefschetz principle was also on the special list.
Let me conclude by quoting Varadarajan's introduction to Harish-Chandra's Collected Papers. Harish-Chandra "survives in his work, which is a faithful reflection of his personality—lofty, intense, uncompromising."

2.7 Les fonctions spéciales : des harmoniques sphériques au programme de Langlands.

Les fonctions spéciales remplissent depuis des siècles les grimoires de formules ; leur abondance est essentielle au travail des mathématiciens, des physiciens, des ingénieurs. Lie, Cartan et Klein avaient signalé que leurs propriétés de symétrie avaient à voir avec les groupes, mais depuis 1927, à mesure qu'on prend conscience du rôle des symétries pour contraindre les équations de la physique, on voit peu à peu les vieilles fonctions spéciales sortir de la théorie des groupes elle-même :

(a) Les harmoniques sphériques, qui servent à la physique depuis Laplace, apparaissent dès qu'on étudie les représentations du groupe des rotations (Hermann Weyl le dit en 1927 : les harmoniques sphériques donnent des bases des espaces de représentations irréductibles).

(b) La fonction Gamma, celle d'Euler, apparaît dès qu'on étudie les représentations unitaires irréductibles du groupe des bijections affines de $\mathbb{R}$ (ce qu'ont fait Gelfand et Naimark en 1946 : elle donne les noyaux des opérateurs intégraux qui font agir le groupe sur des espaces de fonctions).

(c) Les fonctions de Bessel (de première espèce et de paramètre réel) apparaissent dès qu'on étudie les représentations unitaires irréductibles du groupe des déplacements du plan (lorsqu'on réalise une représentation irréductibles comme espace de fonctions sur le plan euclidien, l'unique élément invariant par les rotations autour de l'origine est une fonction de Bessel).

(d) Les polynômes de Gegenbauer, de Legendre et de Jacobi apparaissent dès qu'on étudie les représentations irréductibles du groupe des rotations de $\mathbb{R}^3$, celles de son revêtement double $SU(2)$, et les représentations irréductibles de dimension finie de $SL_2(\mathbb{R})$: ils en donnent les éléments matriciels.

(e) Les polynômes d'Hermite et de Laguerre donnent les éléments matriciels des représentations irréductibles du groupe de Heisenberg.

Notre génération n'a pas connu les encyclopédies de fonctions, pleines de formules mystérieuses qui disent ce qui se passe lorsqu'on ajoute, qu'on multiplie les fonctions spéciales ou lorsqu'on qu'on intègre leurs produits, qui furent longtemps indispensables pour les applications. Mais que la structure de groupe permette, dans *tous* les cas que j'ai mentionnés, de comprendre d'où viennent les propriétés qui alourdissent les bibliothèques, c'est tout à fait remarquable ! En 1955, Wigner en fait le sujet d'un cours à Princeton. Puis la nouvelle traverse l'Atlantique, et de nombreuses fonctions spéciales (je ne vais pas détailler) allongent la liste ci-dessus. En 1968, c'est déjà devenu une industrie, et de nouvelles formules suivent nombreuses ; trois livres paraissent pour l'expliquer − Talman transcrit le cours de Wigner de 1955, Miller et Vilenkin publient leurs études.

Cette industrie est longtemps restée productive, le volume de l'ouvrage de Vilenkin et Klimyk [62] en témoigne. C'est dans deux des ouvrages de 1968 que j'irai chercher au chapitre 6 des candidats pour les champs récepteurs du cervelet vestibulaire.

La théorie des représentations met de l'ordre dans la liste et les propriétés des fonctions spéciales ;
lorsqu'on connaît les invariances géométriques d'un problème,
elle fournit des suggestions de fonctions spéciales à utiliser.

⋆

Puisque j'ai voulu que cette introduction montre comment la généralité grandissante de la théorie des représentations lui permettait d'aborder des problèmes pratiques de plus en plus divers, j'aimerais rappeler brièvement un développement de cette idée en théorie des nombres (voir [21]).

J'ai rappelé plus haut que l'espace des formes automorphes de degré k sur le plan hyperbolique est lié à une représentation de série discrète de $SL_2(\mathbb{R})$. La théorie des nombres connaît bien l'importance de certaines de ces formes automorphes, les *séries théta* qu'on peut attacher aux formes quadratiques, les mêmes dont j'ai dit qu'elles occupaient Frobenius avant 1896. André Weil a beaucoup étudié les travaux de Siegel sur les formes quadratiques et les séries théta, il sait l'influence mystérieuse qu'a le groupe symplectique sur beaucoup de ses résultats. Un jour de 1962, de la théorie quantique des champs, viennent (Shale [56]) un revêtement double du groupe symplectique et une représentation de ce revêtement, et Weil comprend en 1964 qu'une fois ce groupe "métaplectique" adélisé, la représentation (aujourd'hui dite de Shale-Weil) abrite ce qu'il faut pour illuminer les constructions des fonctions théta [65].

Weil ne connaît pas les détails du travail d'Harish-Chandra [46], mais le jeune Langlands se voit offrir une bonne occasion de les connaître : vers 1963 il cherche à calculer les dimensions d'espaces de formes automorphes, et sait que la formule des traces de Selberg ramène cela à un calcul d'intégrale explicite ; il en parle à un jeune collègue qui lui rappelle que ces espaces ont à voir avec

46. C'est Langlands qui le dit ; c'est peut-être étonnant, puisque Weil et Harish-Chandra semblent s'être liés d'amitié.

les représentations de groupes semi-simples... Langlands observe : Harish-Chandra a déjà fait le calcul dont il a besoin, l'intégrale qu'il cherche est celle d'un élément matriciel de représentation de série discrète ! C'est ainsi qu'il commence à concevoir qu'il soit possible d'utiliser les représentations de dimension infinie de groupes réductifs pour éclairer la théorie des nombres.

Vaste programme ! Si vaste qu'une fois incorporée l'approche de Weil et bien d'autres ingrédients qui m'échappent, le tissu de conjectures qui en sortit est plus que célèbre aujourd'hui. Parmi les conjectures de Langlands, certaines concernent directement les fonctions spéciales, par exemple :

> *Les fonctions L de la théorie des nombres, celles qui sont attachées aux extensions galoisiennes de corps de nombres, sont exactement celles qu'une procédure canonique attache à certaines représentations irréductibles de groupes réductifs*[47].

2.8 Les contractions de groupes et de représentations.

Les équations d'onde qui servent à décrire les particules élémentaires sont, d'après Wigner, celles qui sont invariantes par le groupe de Poincaré et dont les espaces de solutions sont irréductibles sous l'action du même groupe. Mais l'équation de Schrödinger n'en fait pas partie : le groupe de Poincaré n'en permute pas les solutions. Elle n'est pas adaptée aux invariances de la *relativité restreinte*, mais à la relativité *galiléenne*. Après ses études sur les représentations du groupe de Lorentz, Bargmann se tourne vers l'équation de Schrödinger, et cherche à comprendre son rapport avec le groupe des changements de référentiels galiléens, le *groupe de Galilée* (voir le chapitre 5). Il comprend qu'elle définit des représentations *projectives* de ce groupe, et que la raison pour laquelle la masse a un statut différent en relativité restreinte et en relativité galiléenne, c'est qu'elle apparaît comme un paramètre déterminant une représentation unitaire dans le travail de Wigner, alors que pour l'équation de Schrödinger, elle est liée au facteur de *phase* qui apparaît lorsque le groupe de Galilée permute les solutions[48] ; Bargmann le dira rapidement à la fin d'une grande étude sur les représentations projectives parue en 1954 [6] − je donnerai une démonstration détaillée de son résultat au chapitre 5.

Au début de l'hiver 1951, Erdal Inonü arrive à Princeton ; il vient de soutenir une thèse à CalTech sur les rayons cosmiques et les mésons π neutres, mais il aime la théorie des groupes et veut passer quelques mois à en faire avant de revenir en Turquie [30]. Il se tourne bien sûr vers Wigner. Ce dernier est au courant du travail que Bargmann est en train de mener sur les représentations projectives ; Wigner suggère à Inonü de compléter ce travail en cherchant si les représentations *unitaires* du groupes de Galilée peuvent avoir aussi un intérêt pour la physique.

Je rappellerai au chapitre 5 leur réponse négative et leurs résultats. Laissons Inonü raconter comment le fait qu'il ne soit pas possible de trouver un intérêt pour la physique aux représentations du groupe de Galilée a mené à la notion mathématique sur laquelle repose la troisième partie de cette thèse (la citation est extraite de [30]).

> *When I reached this point, the original programme proposed to me by Wigner was completed and I started to write the paper on the Galilei representations. But a question remained : how is it that the true representations of the Poincaré group have a physical meaning while those of the Galilei group do not ? Or, in other words, how does the physical meaning disappear when one goes over from the Poincaré group to the Galilei group ?*
> *We thought that at least a partial answer could be obtained by looking at the limits for infinite light velocity of the specific unitary representations of the Poincaré group obtained by Wigner. The idea was to add an appendix to our Galilei paper, giving the results of this limiting process.*

---

47. Ces groupes réductifs ne sont pas réels, bien sûr, ce sont des versions adéliques de groupes réductifs réels.

48. L'aphorisme bien ultérieur de Souriau sur l'analogue de ce fait en mécanique classique ("la masse est la classe de cohomologie du défaut d'équivariance de l'application moment") semble paradoxalement plus célèbre aujourd'hui chez les mathématiciens que le travail de Bargmann.

However, when I tried to take the limits of the unitary representations of the Poincaré group, the outcome became incomprehensible. The limiting process gave a finite answer in some cases, but vanished altogether in other cases.
After we struggled for a couple of weeks without obtaining consistent results, Wigner had the bright idea of separating the problem into its essential components. He said : "Let us first look at the limit of the group, understand what happens there, and then consider the limits of the representations."

Voici le sens qu'ils ont donné à la *"limite pour le groupe"*, avec les mots d'aujourd'hui.

Lorsque G est un groupe de Lie et K est un sous-groupe fermé de G, la contraction G_0 de la paire (G, K) est le produit semi-direct $K \ltimes (\mathfrak{g}/\mathfrak{k})$ associé à l'action (adjointe) de K sur $\mathfrak{g}/\mathfrak{k}$.

Les symboles $\mathfrak{g}$ et $\mathfrak{k}$ désignent bien sûr les algèbres de Lie de G et de K. Je rappellerai au début du chapitre 7 comment la structure de G_0 s'obtient à partir de la structure de G en faisant tendre un paramètre vers zéro dans les relations qui définissent le crochet de Lie de $\mathfrak{g}$; ce paramètre était l'inverse de la vitesse de la lumière dans le travail d'Inonu et Wigner.

L'idée fait écho à des constructions alors récentes de Segal [57]. D'autres auteurs s'intéressent ensuite à la notion de contraction : la thèse de Saletan [54], dirigée par Bargmann, propose notamment des définitions alternatives en 1961. La notion reçoit un examen détaillé dans plusieurs cas d'importance en physique : le groupe de Poincaré est une contraction des groupes de de Sitter, et il a d'autres contractions que le groupe de Galilée (voir Lévy-Leblond [37]). Des études font régulièrement le point sur le sujet, par exemple celle de Monique Lévy-Nahas [38].

La *"limite des représentations"* dont parle Wigner dans la conversation rapportée ci-dessus, en revanche, est plus mystérieuse. Inonü et Wigner définissent deux types de "limites" possibles, dont l'un tient compte du fait que la dimension doit croître au cours de la contraction lorsque G est compact pour espérer obtenir une représentation de G_0. Plusieurs autres cas particuliers sont étudiés par la suite − Mukunda [44] étudie la contraction des représentations de $SL_3(\mathbb{C})$; Mickelsson et Niederle celles du groupe de de Sitter [43]. Plus tard, on verra l'intérêt de la notion pour dire les relations entre fonctions spéciales [13].

Une question revient régulièrement, sans recevoir de réponse générale. Les représentations de la contraction de la paire (G, K) peuvent être déterminées par la méthode de Mackey, et elles sont donc simples à décrire. Quelles sont leurs relations avec celles de G, qui sont en général difficiles à décrire ?

La question est idiote lorsque K est trivial ou est G tout entier ; elle l'est moins, bien sûr, lorsque G est le groupe de Poincaré et que K est le sous-groupe des transformations qui fixent l'axe des temps ; dans ce cas G_0 est le groupe de Galilée et la question est celle qu'Inönü se posait ci-dessus.

En 1975, Mackey lui-même remarque − après avoir observé quelques exemples − que lorsque G est un groupe *semi-simple* et K est un *compact maximal* de G, la question mérite probablement un examen approfondi : voilà déjà vingt ans qu'Harish-Chandra travaille sur le cas où G est semi-simple, et au moins pour ce qui concerne la série principale, ses résultats semblent reflétés, bien que ce soit de façon appauvrie, par la structure des représentations de G_0. Des considérations de mécanique quantique (rappellées en 3.3 ci-dessous) semblent indiquer que c'est moins surprenant que la structure algébrique de G et celle de G_0 ne le laisseraient penser. Peut-on aller plus loin ?

C'est à la question de Mackey que la troisième partie de cette thèse essaie de répondre ; je donnerai plus de détails sur les idées de Mackey et leur postérité dans la section 3.2 de cette introduction, et bien sûr au chapitre 7.

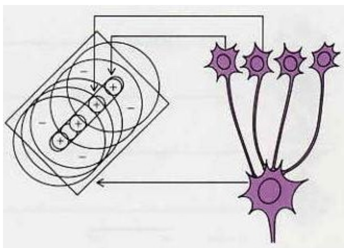
3 Thèmes de cette thèse et résumé de ses résultats.

Il est temps de décrire les résultats de ma thèse. Six des huits chapitres à venir reprennent des textes publiés (chapitres 2 et 3) ou disponibles dans des archives ouvertes (chapitres 1, 4, 7, 8) ; le chapitre 5 peut être lu indépendamment du chapitre 6, qui en dépend largement. Cela a plusieurs conséquences sur l'organisation du manuscrit :

– Chaque chapitre (excepté le sixième) a été rédigé pour être lu indépendamment des autres, et comporte un résumé, une introduction indépendante et sa propre bibliographie. Même si certains des points de vue énoncés dans cette introduction générale n'apparaîtront plus dans le texte des chapitres, c'est donc la redondance qui sera la règle tout au long du manuscrit. On trouvera les mêmes textes cités de nombreuses fois, des figures reproduites à l'identique dans des chapitres distincts, etc.
– Dans la mesure où plusieurs thèmes ou objets techniques reviennent régulièrement même dans des contextes éloignés, plusieurs définitions ou théorèmes classiques seront rappelés plusieurs fois dans des termes très proches. Les parties I et II reviennent plusieurs fois sur des constructions de base, notamment lorsqu'elles concernent les probabilités où je suis particulièrement ignorant. J'espère pouvoir compter sur votre patience quand vous lirez pour la quatrième fois la définition des ondes d'Helgason, ou pour la troisième fois le rappel des propriétés de base des champs gaussiens ou des fonctions sphériques.
– Les chapitres 2 et 3 ont paru dans une revue qui n'est pas destinée aux seuls mathématiciens, et le chapitre 6 présente un récit de travail expérimental manqué plutôt qu'un ensemble de théorèmes ; seuls les chapitres 1, 4, 7 et 8 sont donc sous une forme mathématique tout à fait traditionnelle. La plupart des résultats des chapitres 2 et 3 peuvent par ailleurs, si on oublie les discussions qui se réfèrent au cerveau, être vus comme des cas particuliers bidimensionnels des résultats du chapitre 4 ; la longueur de la première partie est largement due à ce fait.

3.1 Première partie : études motivées par le cortex visuel primaire

3.1.a Effet de moiré sur les espaces symétriques de type non-compact



Pour comprendre comment les profils récepteurs des neurones de V1 pouvaient avoir des préférences directionnelles alors que les neurones du corps genouillé latéral qui sont situés en amont n'en ont pas, Hubel et Wiesel ont proposé un scénario célèbre : il suffit qu'un neurone de V1 reçoive des informations de neurones géniculaires dont les centres des profils récepteurs sont alignés et voisins, puis *somme* les activités de ces neurones.

Dans ce court chapitre rédigé en juin 2014 (version française légèrement abrégée sur arXiv :1602.03871), je pars d'un constat simple de l'analyse de Fourier classique :

> Sur le plan euclidien,
> l'interférence constructive de fonctions de Bessel
> dont les centres de symétrie sont disposés le long d'une droite
> permet de reconstruire une onde plane.

J'ai dit plus haut que les fonctions de Bessel avaient une relation toute spéciale aux représentations du groupes des déplacements du plan : ce sont les seuls éléments invariants par rotation dans les espaces de fonctions qui donnent les composantes irréductibles de l'espace des fonctions sur $\mathbb{R}^2$.

En exprimant la préférence directionnelle à l'aide de l'invariance par les translations selon une droite (et en remplaçant les filtres de Gabor par des ondes planes, sacrifiant le réalisme à l'expression des symétries), le constat ci-dessus permet de mettre au jour comment l'idée d'Hubel et Wiesel dépend de la structure du groupe des déplacements.

Si le plan visuel était hyperbolique, le scénario d'Hubel et Wiesel aurait-il un analogue qui pourrait se comprendre à l'aide des représentations du groupe d'isométries du plan hyperbolique, qui est isomorphe à $SL_2(\mathbb{R})$? Je montre ici que c'est le cas, et qu'un résultat analogue existe pour tous les espaces riemanniens symétriques de courbure négative.

Soit G un groupe semi-simple réel non compact et K un compact maximal de G. L'analogue d'une fonction de Bessel pour l'espace symétrique G/K, c'est le seul élément K-invariant d'un facteur irréductible de $\mathbf{L}^2(G/K)$, une *fonction sphérique d'Harish-Chandra*.

Les *horocycles* de G/K sont des sous-variétés qui jouent un rôle important pour l'analyse harmonique sur G/K ; si G est $SU(1,1)$ et si G/K est vu comme le disque de Poincaré, ce sont les cercles tangents au bord du disque. Les horocycles viennent en famille, par exemple sur le disque la famille des cercles tangents au même point. Dans le cas général, chaque famille d'horocycles est l'ensemble des orbites dans G/K d'un conjugué du groupe[49] N qui apparaît dans la décomposition d'Iwasawa.

Harish-Chandra a montré qu'il y existe des fonctions spéciales sur G/K, dont chacune est constante sur une famille d'horocycles, qui permettent de construire les fonctions sphériques par interférence constructive. Sigurdur Helgason a compris que ces fonctions spéciales sont des *ondes*, et peuvent jouer pour l'analyse harmonique sur G/K le rôle que jouent les ondes planes pour l'analyse de Fourier classique.

Il n'est pas difficile de montrer le résultat suivant, qui est l'objet du chapitre 1 :

Sur un espace symétrique de type non-compact,
l'interférence constructive de fonctions sphériques d'Harish-Chandra
dont les centres de symétrie sont disposés le long d'un horocycle
permet de reconstruire une onde d'Helgason.

3.1.b Cartes d'orientation de V1 et champs aléatoires gaussiens à symétrie euclidienne

Le chapitre précédent partait des profils récepteurs des neurones individuels − la question **(a)** de la section 1.1. Dans la suite de la première partie, je me tourne vers les cartes d'orientation de V1 (et donc vers la question **(b)**) pour faire le point sur le rôle des arguments de symétrie dans un modèle dû à Fred Wolf et Theo Geisel (1998), qui fut le premier à prédire la densité π des pinwheels. Ce modèle est aussi le point de départ des chapitres 3 et 4.

Présentation du premier modèle de Wolf et Geisel. Soit $\mathcal{C}$ la région centrale de la surface de V1. Le modèle traite la carte des orientations comme si elle était issue d'une fonction à valeurs *complexes* sur $\mathcal{C}$, en en prenant l'argument ou le demi-argument ; les pinwheels correspondent alors aux zéros de cette fonction.

Le modèle est probabiliste : il propose de voir chaque carte comme un *tirage typique* d'une variable aléatoire, et fait donc intervenir une fonction aléatoire (le terme consacré est *champ aléatoire*) sur $\mathbb{R}^2$ à valeurs complexes, disons $\mathbf{z}$. Il y a au moins deux bonnes raisons de le faire : la première est que les cartes d'orientation observées chez deux individus différents ne sont en général pas superposables et qu'il est donc utile de chercher à décrire

49. Qui n'est pas unique, mais qui l'est à conjugaison près.

la *loi* du champ plutôt que chaque réalisation individuelle, et la seconde (très importante pour Wolf et Geisel) est qu'on ne connaît pas l'état initial de la carte, et qu'il peut faire la part belle à des fluctuations aléatoires survenant au début du développement.

Il est biologiquement raisonnable de tenter, au moins pour ce qui concerne les premiers stades du développement, d'obtenir des résultats réalistes à partir d'un champ aléatoire *gaussien* sur $\mathcal{C}$. Mathématiquement, c'est presque indispensable, puisqu'il y a peu de résultats généraux disponibles dans un contexte proche du nôtre lorsqu'on sort de cette classe de processus aléatoires.

Wolf et Geisel ajoutent, de façon cruciale, une hypothèse de symétrie : ils *assimilent $\mathcal{C}$ à un plan euclidien et supposent que la loi du champ est invariante par les rotations et les translations de $\mathcal{C}$.*

Ils ajoutent enfin une hypothèse de *finesse spectrale* pour rendre compte de la quasipériodicité observée sur les cartes : ils supposent que presque tous les tirages de $\mathbf{z}$ ont leur transformée de Fourier concentrée sur un même cercle. On dit alors que la carte est *monochromatique*, pour signifier que le spectre des tirages de $\mathbf{z}$ ne comporte qu'une longueur d'onde (le terme ne renvoie donc ni à la couleur utilisée pour représenter la carte, ni à la fréquence spatiale préférée de chaque neurone individuel).

Wolf et Geisel démontrent alors le résultat suivant, découvert indépendamment et à peu près simultanément par Berry et Dennis en optique (dans des travaux sur les superpositions aléatoires d'ondes lumineuses de même fréquence) :

Le nombre moyen de zéros d'un champ aléatoire gaussien sur $\mathbb{R}^2$,
lorsqu'il est invariant par rotations et translations et monochromatique,
est π
dans chaque région dont l'aire est le carré de la longueur d'onde caractéristique.

Pour Wolf et Geisel, une région qui a cette aire mérite d'être assimilée à une hypercolonne.

⋆

J'insisterai au chapitre 3 sur le fait que la condition de finesse spectrale signifie que les tirages de $\mathbf{z}$ se concentrent sur un facteur irréductible de la décomposition de $\mathbf{L}^2(\mathbb{R}^2)$ en intégrale directe de représentations irréductibles du groupe des déplacements. Arrêtons-nous pour résumer le point de départ des chapitres 2, 3 et 4 qu'on va lire :

Certaines des propriétés importantes des cartes d'orientation de V1
sont plutôt bien reproduites par les tirages "typiques" d'un champ aléatoire gaussien
dont la loi est invariante par le groupe des déplacements
et dont les tirages explorent un facteur irréductible de la représentation quasi-régulière.

Ce modèle n'est qu'un premier essai ; Wolf et Geisel eux-mêmes insistent depuis plus de quinze ans sur le fait qu'il y en a de bien meilleurs pour parler des cartes d'orientation à un stade plus avancé du développement cortical, et ils travaillent avec leurs collaborateurs à en trouver. Celui qui a été proposé en 2010 dans l'appendice de l'article de Kaschube et al. qui présente les résultats expérimentaux sur la densité π (voir la fin du chapitre 3) donne d'excellents résultats numériques.
Mais les traits mathématiques des champs gaussiens qu'ils utilisent sont remarquables, surtout compte tenu du lien entre champs gaussiens invariants et éléments matriciels de représentations rappelé au début du chapitre 4 ; leur résultat mathématique sur la densité π mérite d'être étudié et généralisé – c'est ce que je fais dans les chapitres 2, 3 et 4 de cette thèse.

Contenu du chapitre 2. Il s'agit d'un assez court article paru [1] au Journal of Mathematical Neuroscience en avril 2015. Le mot "groupe" n'y apparaît pas en toutes lettres.

Son but est de faire le point sur le rôle de l'hypothèse de finesse spectrale qui est si remarquable du point de vue de la théorie des groupes, et de savoir si cette hypothèse est essentielle pour obtenir la densité π. Contrairement à ce qu'on pourrait penser à première vue, l'hypothèse de finesse spectrale n'est pas très réaliste : je dirai comment les données expérimentales récentes de la thèse de Schnabel tendent plutôt à montrer que le pic dans le spectre des corrélations de chaque carte individuelle n'est *pas* très fin.

Je commence d'abord par préciser ce que doit être la "taille des hypercolonnes" dans un champ gaussien sur $\mathbb{R}^2$ (à valeurs complexes et à symétrie euclidienne) lorsqu'on retire l'hypothèse de finesse spectrale, en observant la distance à laquelle on peut espérer voir se reproduire la spécialité d'orientation observée en chaque point.

> L'espacement caractéristique des cartes
> issues d'un champ gaussien à symétrie euclidienne
> est donné par la longueur d'onde associée au nombre d'onde quadratique moyen
> du spectre de puissance du champ.

En observant la démonstration du résultat de Wolf-Geisel et Berry-Dennis, dans la version mathématiquement complète précisée récemment par Jean-Marc Azaïs et Mario Wschebor (voir aussi le livre d'Adler et Taylor), je constate le fait suivant.

> La densité de pinwheels des cartes issues d'un champ gaussien,
> exprimée dans l'unité d'aire donnée par l'espacement caractéristique,
> vaut π, même si le spectre de puissance ne se concentre pas sur un facteur irréductible.

Cela montre que le rôle de la condition de finesse spectrale n'est *pas* de garantir une densité π. Il est en revanche difficile de ne pas croire que cette condition n'est pas essentielle pour observer des cartes "quasipériodiques", même si le sens à donner à "quasipériodique" est vague. Comment préciser cela ?

Je me tourne pour le faire vers la *variance* des espacements constatés entre neurones qui partagent la même spécialité d'orientation. Il est naturel de penser que cette variance est minimale lorsque le champ est monochromatique, mais compte tenu de l'importance apparente de la structure des cartes pour le traitement de l'information par l'aire V1, peut-on s'en assurer rigoureusement ?

Des travaux anciens de Cramer et Leadbetter (1967) permettent d'écrire une formule close pour la quantité qui permet de mesurer cette variance, mais qu'elle est si compliquée qu'il semble impossible d'en extraire quelque renseignement que ce soit. J'ai donc eu recours au calcul numérique (qui ne va pas de soi pour une expression de ce type, où il y a des intégrales oscillantes à évaluer et, pire, à dériver par rapport à des paramètres) pour m'assurer du fait suivant.

> La variance des espacements locaux dans les cartes issues d'un champ gaussien
> est minimale lorsque le spectre de puissance du champ se concentre sur un facteur
> irréductible.

3.1.c Chapitre 3 : cartes d'orientation sur des espaces non-euclidiens

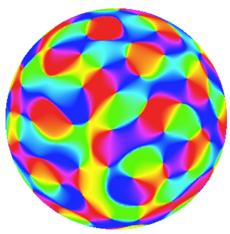
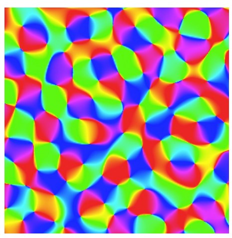
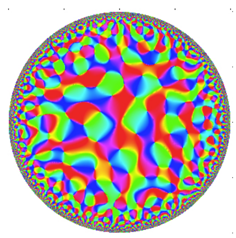
Ce chapitre est un article, plus long que le précédent, paru au Journal of Mathematical Neuroscience en juin 2015 [2]. Son but est de signaler l'interprétation du modèle de Wolf et Geisel au moyen de la théorie des représentations (voir le second encadré bleu du paragraphe 3.1.b) et de se servir de cette interprétation pour obtenir des structures en pinwheels sur la sphère et le plan hyperbolique à l'aide de la théorie des représentations de leurs groupes d'isométrie.

Assimiler la région centrale de la surface de V1 à un plan euclidien est utile, et c'est même naturel : la région centrale de V1 est à peu près plate chez le tupaya qui est l'une des espèces de références pour les expériences sur V1, et les techniques d'imagerie optique utilisées de nos jours enregistrent plutôt la projection de la surface corticale sur le plan focal des capteurs optiques. Cela dit, la surface de V1 peut être courbe (elle l'est chez l'humain), et surtout, il est important de distinguer entre la géométrie *du plan visuel* où Poincaré et Souriau nous signalaient en 1.4 qu'il est essentiel que le cerveau sache identifier des déplacements rigides et des lignes droites [50], et la géométrie *du plan cortical* qui est le cadre de l'action de groupe utilisée dans les modèles qui cherchent à décrire V1. Vu les propriétés de l'application rétinotopique présentée au paragraphe 1.2.b.1, une translation dans le plan visuel ne se traduit *pas* par une translation dans le plan cortical, et une rotation dans le plan visuel change nécessairement la distribution des activations des neurones individuels à la surface de V1, sans qu'il soit facile de savoir si elle induit une rotation dans cette distribution des activations (des résultats remarquables de Bosking [11], faisant écho à des idées de Field et Hayes sur le "champ d'association" qui nous permettrait de reconnaître des contours, vont dans ce sens − voir le chapitre 4 dans le livre de Petitot, et [69] − mais semblent faire débat).

Est-il essentiel que la structure du groupe d'isométries *de la surface du cortex* soit celle du groupe des déplacements, dont Poincaré signalait le caractère remarquable, pour obtenir les structures en pinwheels de V1 au moyen de la théorie des représentations et la densité π ? Ce chapitre montre que ce n'est pas le cas.

Je signale qu'on construit facilement de telles structures sur la sphère et sur le plan hyperbolique à l'aide de résultats anciens de Yaglom (1961) sur les champs gaussiens sur les espaces homogènes, lorsqu'on y ajoute les constructions explicites dues à Weyl et Élie Cartan dans le cas sphérique et à Helgason dans le cas hyperbolique.

Sur le plan hyperbolique et sur la sphère,
les champs aléatoires gaussiens à valeurs complexes dont la loi est invariante
et dont les tirages explorent un facteur irréductible de la représentation quasi-régulière
fournissent des structures en pinwheels.

Le résultat d'Azaïs et Wschebor utilisé au chapitre 2 fait partie d'un vaste ensemble de généralisations récentes de la formule de Kac-Rice (qui étudie les zéros d'une fonction aléatoire depuis que Rice s'est saisi de ces objets mathématiques pour étudier le bruit dans les conversations téléphoniques en 1944). J'utilise cette formule classique pour estimer la distance géodésique moyenne entre deux points de la sphère ou du plan hyperbolique où on trouve la même "orientation" dans les cartes analogues à celles de V1 que je viens de construire, et cela fait apparaître le spectre du laplacien sur chacun de ces deux espaces. Je fais alors appel aux résultats d'Azaïs et Wschebor pour montrer que le résultat dû à

50. Le chapitre 1 consistait à changer de groupe d'isométries *pour cette géométrie-là*, contrairement à ce que nous allons faire ici.

Wolf-Geisel et Berry-Dennis dans le cas euclidien monochromatique s'étend en fait aux deux espaces symétriques envisagés dans ce chapitre, sans restriction sur la courbure ni sur la finesse spectrale du champ.

Sur le plan hyperbolique et sur la sphère,
les champs aléatoires gaussiens dont la loi est invariante
ont une densité de zéros égale à π.

3.1.d Chapitre 4 : champs gaussiens invariants sur les espaces homogènes

Ce chapitre, rédigé en août 2015 (arXiv :1602.02560), contient la généralisation des résultats du chapitre 2 aux espaces homogènes *de dimension arbitraire*. En dimension deux, il n'y a que trois types d'espaces symétriques, et leur géométrie est très célèbre. Mais les résultats de Yaglom que je signalais au chapitre précédents sont très généraux ; ils répondent partiellement à une question posée très tôt (1944) par Kolmogorov, qui pensait que les champs aléatoires invariants par une action de groupe trouveraient de nombreuses applications pratiques.

En 1961, nous avons vu que la théorie des représentations de groupes non-compacts était en plein développement. Yaglom signale plusieurs résultats, dus notamment à Gelfand, qui mènent en principe à des constructions explicites. Plus de cinquante ans après, nous disposons bien sûr d'outils que n'avait pas Yaglom, et il m'a paru souhaitable de faire le point sur ses résultats en les rendant aussi explicites que possible, afin de pouvoir par exemple simuler des tirages de champs gaussiens à l'aide d'un ordinateur. Je commence donc par passer en revue quelques constructions de fonctions sphériques élémentaires sur les espaces homogènes (ces fonctions classifient les champs gaussiens invariants à égalité en loi près), et décris les champs gaussiens associés.

Sur de nombreux espaces homogènes provenant de paires de Gelfand,
les champs aléatoires gaussiens à valeurs complexes dont la loi est invariante
et dont les tirages explorent un facteur irréductible de la représentation quasi-régulière
admettent des descriptions explicites.

Je démontre ensuite que le résultat constaté par Berry-Dennis et Wolf-Geisel sur les champs monochromatiques à symétrie euclidienne en dimension 2, qui était si surprenant du point de vue de la physique ou de la biologie, est en fait le cas le plus simple d'un phénomène qui concerne *tous* les champs gaussiens invariants (à valeurs dans un espace vectoriel de dimension finie, mais arbitraire) sur les espaces homogènes riemanniens.

Sur un espace homogène riemannien X,
la densité de zéros d'un champ gaussien à valeurs dans V dont la loi est invariante
lorsqu'elle est exprimée dans l'unité de volume caractéristique du champ,
ne dépend que des dimensions de X et V, et pas du groupe qui agit sur X.
La valeur de la densité est $\frac{(\dim X)!}{(\dim X - \dim V)!}\left(\frac{\pi}{2}\right)^{\dim X/2}$.

Les résultats du chapitre 3 laissaient penser que l'analyse non-commutative est importante pour *construire* explicitement les champs gaussiens invariants, mais que leur densité de zéros, une fois exprimée dans une unité de volume adéquate (reliée au spectre du laplacien et au spectre de puissance du champ), ne dépend pas de la structure du groupe. Ce que je viens d'encadrer montre que ce fait n'est pas spécifique à la dimension 2. Pour que l'intérêt de ce résultat soit clair et pour que la définition de l'unité de volume appropriée ne semble pas étrange, il est probablement utile que Berry-Dennis et Wolf-Geisel aient découvert le phénomène dans un contexte où il y avait une unité de volume privilégiée "évidente".

3.2 Deuxième partie : représentations du groupe de Galilée ; recherche de profils récepteurs de neurones vestibulaires

Cette partie contient peu de résultats nouveaux, et son esprit est inhabituel. J'ai eu au cours de ma thèse une très belle occasion de tenter d'utiliser la théorie des représentations du groupe de Galilée pour étudier des données expérimentales sur la région du cervelet qui participe au traitement de l'information vestibulaire : des enregistrements remarquables de l'activité électrique de neurones de cette région venaient d'être réalisés à l'Institut de Biologie de l'Ecole Normale Supérieure dans le nodulus de rats vigiles et se déplaçant librement. Matthieu Tihy, Guillaume Dugué et Clément Léna nous ont proposé de reprendre l'idée de Daniel Bennequin et Alain Berthoz [9], enthousiasmante *a priori* pour les neurosciences, selon laquelle les éléments matriciels de représentations du groupe de Galilée étaient susceptibles d'être utiles pour décrire les résultats de ces enregistrements.

Je rappelle dans le chapitre 5 que cette idée est tout à fait testable en principe, grâce à la connaissance explicite des éléments matriciels obtenue dans les traités de 1968 que je mentionnais au paragraphe 2.7. La forme numérique des données obtenues par Matthieu Tihy était tout à fait adaptée pour le faire, et j'ai passé un temps certain à explorer numériquement notre hypothèse. Le chapitre 6 est le récit de ce travail. Hélas ! Bien que quelques résultats positifs soient sortis de nos analyses, aucun ne nécessitait en dernière analyse de recourir à la théorie des représentations.

Je pense cependant que l'énoncé de notre hypothèse et le récit de cette collaboration ont leur place dans ma thèse : l'idée qui ne s'est pas avérée fructueuse pour l'étude du vestibulo-cervelet pourrait bien être pertinente pour d'autres aires vestibulaires, et si c'était le cas, les formules explicites pour les éléments matriciels seraient peut-être utiles. Le chapitre 5 sera par ailleurs l'occasion de faire le point sur certains aspects importants de la théorie des représentations du groupe de Galilée et d'indiquer une démonstration simple du résultat de Mackey (1949) qui est le fondement des chapitres 7 et 8.

3.2.a Chapitre 5 : quelques rappels sur le groupe de Galilée et la théorie de Mackey

Ce chapitre contient essentiellement des rappels destinés à préparer les études des chapitres 6, 7, 8.

Au vingtième siècle, les représentations du groupe de Galilée ont été utiles surtout quand elles sont projectives : j'ai rappelé au paragraphe 2.8 que Bargmann avait signalé dès 1954 que l'espace des solutions de l'équation de Schödinger abrite une représentation projective de ce groupe, et qu'Inonü et Wigner n'avaient pas trouvé d'intérêt pour la physique aux représentations qui sont unitaires.

Je commence par faire le point sur les résultats de Bargmann, Inonü et Wigner, et d'abord par remarquer le fait suivant.

> L'équation de Schrodinger se cache dans la seule structure du groupe de Galilée :
> les résultats de Mackey sur les produits semi-directs l'exhibent naturellement
> dès qu'on étudie les représentations projectives du groupe de Galilée
> dont la classe de cohomologie correspond à une masse fixée.

Comme je ne connais pas d'étude qui signale la possibilité d'obtenir l'équation de Schödinger *a priori* en partant de la structure du groupe plutôt que de partir de l'équation supposée connue et de constater son rôle dans la classification des représentations projectives, j'ai pris la peine de donner une démonstration complète : elle revient à la première étape de la démonstration des résultats de Mackey sur les produits semi-directs. Ces derniers sont la clé de la liste des représentations irréductibles indiquée par Inonü et Wigner et de la troisième partie de cette thèse toute entière, et je saisis l'occasion pour transcrire, d'après Niels Skovhus Poulsen et Bent Ørsted, une démonstration qui me semble, pour les cas (groupe de Galilée, groupes de déplacements de Cartan) qui m'occupent dans cette thèse, plus facilement (ou au moins plus rapidement) lisible que les rédactions générales des excellentes références [41, 60, 32].

Une fois décrite la classification des représentations unitaires irréductibles d'après d'Inonü et Wigner, je rappelle les formules obtenues en 1968 par Vilenkin (cas des rotations) et Willard Miller (cas des transformations galiléennes qui ne changent pas l'origine de l'espace-temps) pour les éléments matriciels des représentations unitaires irréductibles de la partie homogène du groupe de Galilée :

Les éléments matriciels des représentations unitaires du groupe des déplacements de $\mathbb{R}^3$
(et donc de la partie linéaire du groupe de Gailiée)
peuvent être écrits sous forme close et évalués numériquement.

3.2.b Chapitre 6 : récit d'une étude sur l'activité de neurones du cervelet vestibulaire

Le vestibulo-cervelet est l'une des structures essentielles pour le traitement de l'information issue des canaux semi-circulaires et des otolithes. J'expliquerai pourquoi, compte tenu de ce que j'ai rappelé aux paragraphes 1.3 et 1.4, il est raisonnable de chercher des *profils récepteurs* pour les neurones (cellules de Purkinje) du vestibulo-cervelet à l'aide des éléments matriciels identifiés au chapitre précédent.

Ce chapitre est le récit de notre tentative en ce sens, en collaboration avec Matthieu Tihy, Guillaume Dugué, Clément Léna, Boris Gourévitch et Daniel Bennequin. Il ne contient pas de théorème, mais j'y rappelle quelques idées courantes sur le cervelet et les cellules de Purkinje de la région enregistrée à l'IBENS, et j'y raconte comment j'ai exploré les données numériquement (avec le langage Python) et les précautions que j'ai prises pour le faire, puis je présente les conclusions (souvent décevantes) auxquelles nous avons abouti.

Nous obtenons des résultats qui sont tout à fait compatibles avec les plus simples des idées courantes sur le vestibulo-cervelet, par exemple une vérification très claire du fait suivant.

Il y a des neurones dont l'activité électrique est très bien décrite par
une combinaison linéaire des vitesses angulaires de la tête au même instant,
exprimées dans un référentiel lié à la tête selon trois axes orthogonaux..

Mais hélas ! Contrairement à ce que nous espérions, la théorie des représentations ne nous a été d'aucun secours pour décrire l'activité des autres neurones.

Nous n'avons pas réussi à voir de neurones dont l'activité électrique est bien décrite
par une fonction des positions de la tête dans un référentiel allocentrique.

Pour les neurones qui réagissent aux rotations <u>et</u> translations dans un référentiel égocentrique,
nous n'avons pas vu que l'activité soit bien décrite par des coefficients de représentations
du groupe de Galilée homogène.

3.3 Troisième partie : correspondance de Mackey entre les représentations d'un groupe de Lie réductif et celles de son groupe de déplacements de Cartan

Cette partie étudie la contraction de la paire formée par un groupe semi-simple réel non compact et un sous-groupe compact maximal. J'essaie de donner corps à une idée formulée en 1975 par G. W. Mackey, reprise et précisée récemment par Nigel Higson.

Disons rapidement d'où est venue l'idée de Mackey. Soit G le revêtement universel du groupe de Lorentz ; comme je le disais au paragraphe 2.5, il est isomorphe à $SL_2(\mathbb{C})$. Un compact maximal est le sous-groupe (simplement connexe) K qui se projette sur le groupe des rotations spatiales qui fixent l'axe des temps, isomorphe à $SU(2)$. Notons G_0 la contraction de la paire (G, K) : c'est un revêtement double du groupe des déplacements de l'espace euclidien. Comme je l'ai rappelé, sa structure algébrique est bien différente de celle de G.

En mécanique quantique, les manuels nous disent que l'espace des états instantanés d'une particule libre "non relativiste" de spin j (mais pas nécessairement élémentaire ; cela peut être un atome) est contenu dans $\mathbf{L}^2(\mathbb{R}^3, \mathbf{H}_j)$ où $\mathbf{H}_j$ est un espace de dimension finie qui porte la représentation irréductible de $SU(2)$ indiquée par le spin j, disons μ_j. Une variante de l'idée de Wigner nous dit que cet espace porte une représentation de G_0, peut-être irréductible, et pour des raisons de structure, cette représentation est nécessairement contenue dans la représentation induite $\mathcal{R}_0 := Ind_K^{G_0}(\mu_j)$.

Les facteurs irréductibles de la représentation $\mathcal{R}_0$, Mackey les connaît bien : comme je l'ai rappelé plus haut, c'est lui qui a développé la théorie des représentations pour G_0. Il sait qu'ils correspondent à des états de particules dont l'hélicité et l'impulsion totale sont fixées. Dans une conférence de 1971 à Budapest, il commence par remarquer que si notre espace (celui du quotidien) avait été de courbure constante mais légèrement négative, ce qui après tout est imaginable, la même discussion aurait abouti à chercher une sous-représentation de $\mathcal{R} = Ind_K^G(\mu_j)$. La décomposition de $\mathcal{R}$ en facteurs irréductibles peut-elle faire apparaître des paramètres radicalement différents et des structures radicalement différentes ? Mackey dit que l'interprétation physique laisse penser que ce n'est pas le cas ; puis en reprenant Gelfand et Naimark, il constate que les couples (impulsion, hélicité) qui sont apparus dans $\mathcal{R}_0$ permettent bien d'identifier les facteurs irréductibles de $\mathcal{R}$, et que la façon dont Gelfand et Naimark ont construit les représentations de série principale rappelle tout à fait celle dont sont construits les facteurs irréductibles de $\mathcal{R}_0$.

La série principale existe aussi pour tous les groupes semi-simples réels non compacts. Mackey regarde prudemment $SL_2(\mathbb{R})$, $SL_3(\mathbb{R})$, $SL_n(\mathbb{C})$, $SL_n(\mathbb{R})$ et constate qu'au-delà de la série principale, il reste vrai que les paramètres d'Harish-Chandra pour identifier les représentations qui servent à décomposer la représentation quasi-régulière ressemblent beaucoup à ceux qui sortent de la méthode du "petit groupe" pour étudier les représentations du groupe qu'on obtient par contraction.

Par ailleurs, si G est un groupe semi-simple réel non compact et K est un sous-groupe compact maximal, G/K est un espace symétrique de courbure négative (qui peut être rendue aussi petite qu'on le souhaite par un changement d'échelle), alors que G_0/K est un espace euclidien. Élie Cartan avait déjà noté que les espaces symétriques venaient en famille, et G_0 est appelé *groupe de déplacements de Cartan* de G dans ce cas.

On peut donc reprendre la question précédente et se demander s'il y a des analogies entre la théorie des représentations de G et celle de G_0. Dans l'article de 1975 qui fait suite à sa conférence, Mackey demande bravement : se peut-il qu'il y ait une correspondance "naturelle" entre "presque toutes" les représentations de G et "presque toutes" celles de G_0, malgré la différence de structure entre les deux groupes et surtout malgré l'écart entre la difficulté de la théorie des représentations de G et celle de G_0 lorsqu'on s'éloigne de la série principale ?

⋆

Les réactions à cette idée semblent avoir été assez sceptiques. La correspondance proposée par Mackey pour $SL_3(\mathbb{R})$ est assez étrange au premier abord, et lorsque le groupe G admet une série discrète, pour laquelle la construction des espaces de Hilbert était un sujet brûlant, les remarques de Mackey sont pluôt mystérieuses vues de 1971. La notion de K-type minimal, qui me sera essentielle pour définir une correspondance entre le dual tempéré de G et le dual unitaire de G_0 au chapitre 7, attendait la thèse de Vogan (1976) ; la classification complète des représentations irréductibles tempérées de G attendrait le travail de Knapp et Zuckerman de 1982.

Peu d'études reprennent en détail la suggestion de Mackey ; la plus complète jusqu'à 2007 est, à ma connaissance, celle de Dooley et Rice [18] qui démontrent en 1985 que pour chaque représentation π_0 de G_0 correspondant à une orbite *de dimension maximale* de K dans $(\mathfrak{g}/\mathfrak{k})^\star$ (voir le chapitre 5 pour le rapport avec la description de Mackey du dual unitaire de G_0), il existe une famille de représentations *de série principale* dont les opérateurs convergent faiblement vers ceux de π_0 au cours de la contraction.

Cela dit, le lien entre l'analyse harmonique des *fonctions* sur G/K (qui ne nécessite, d'après Harish-Chandra et Helgason, que la série principale) d'une part, et l'analyse de Fourier classique sur l'espace euclidien G_0/K d'autre part, est devenu un thème classique peu après 1975 : Helgason s'est penché de très près sur la question en 1980 [24], et la *conjecture de Kashiwara-Vergne* sur la formule de Campbell-Hausdorff pour un groupe de Lie général, ses liens avec l'isomorphisme de Duflo qui permet de réduire l'étude des opérateurs différentiels bi-invariants sur G à celle des opérateurs différentiels linéaires à coefficients constants sur $\mathfrak{g}$, ont suscité de nombreuses études depuis plus de trente ans (voir la belle synthèse de Rouvière [53], qui contient beaucoup de résultats importants de ce dernier sur le sujet). Ces études sont différentes dans leur esprit de ce que cherchait Mackey (et n'ont pas besoin de le citer).

⋆

Alain Connes reprend la question de Mackey, de façon détournée, dans les années 1980, mais sa réponse (bien que ce soit le résultat le plus général que je connaisse sur le sujet) s'éloigne de la théorie des représentations. Connes remarque que lorsque G est un groupe de Lie connexe et lorsque K en est un compact maximal, bien qu'il soit difficile d'observer *directement* les analogies entre le dual unitaire de G et celui de G_0 comme Mackey l'espérait, il est possible d'établir qu'ils ont une ressemblance *de nature cohomologique* : les groupes de K-théorie de leurs $C^\star$-algèbres de groupe sont isomorphes, et une reformulation due à Connes et Higson de la désormais célèbre conjecture de Baum-Connes fournit un isomorphisme.

C'est encore de la géométrie non-commutative qu'est venue la première étude qui fasse suite à la question de Mackey et concerne le dual tempéré tout entier. En 2008, Nigel Higson se concentre sur le cas des groupes semi-simples *complexes*. En utilisant le fait que le dual tempéré se réduit à la série principale dans le cas des groupes complexes et une coïncidence remarquable et apparemment jamais signalée, la connexité des "petits groupes" de la machine de Mackey de G_0 lorsque G est complexe, il montre qu'il existe en fait une *bijection naturelle* entre le dual tempéré de G et le dual unitaire de G_0.

La bijection découverte par Higson n'est pas un homéomorphisme entre le dual tempéré de G et celui de G_0, mais elle est assez naturelle pour permettre de donner une démonstration très simple de la conjecture de Baum-Connes (déjà connue bien sûr pour ces groupes) où l'isomorphisme entre les groupes de K-théorie des $C^\star$-algèbres réduites de G et G_0 apparaît comme un simple reflet des ressemblances entre les $C^\star$-algèbres réduites elles-mêmes que révèle la correspondance de Mackey.

Higson signale aussi que dans la mesure où la conjecture de Baum-Connes explore la structure topologique du dual tempéré plutôt que sa structure d'espace mesuré chère à Mackey, et où elle est démontrée pour tous les groupes de Lie, il est probable qu'une telle bijection existe aussi pour les groupes semi-simples réels.

3.3.a Paramétrisation commune du dual tempéré du groupe semisimple et du dual unitaire du groupe de Cartan

Le chapitre 7, rédigé en juin 2015 (arXiv :1510.02650), est l'un des plus longs de cette thèse ; c'est celui qui m'a demandé le plus de travail.

Je commence par y décrire une bijection naturelle entre le dual tempéré d'un groupe de Lie semi-simple et le dual unitaire de son groupe de déplacements de Cartan.

Soit G un groupe de Lie réductif réel, supposé linéaire, connexe et pour l'instant de centre compact, et soit K un compact maximal de G. En reprenant les notations $\mathfrak{g}$ et $\mathfrak{k}$ ci-dessus, notons $\mathfrak{p}$ l'orthogonal de $\mathfrak{k}$ dans $\mathfrak{g}$, et $\mathfrak{a}$ une sous-algèbre abélienne maximale de $\mathfrak{g}$ contenue dans $\mathfrak{p}$. D'après les résultats de Mackey rappelés au chapitre 5 et les propriétés de structure habituelles des groupes de Lie semisimples, pour construire une représentation de G_0, il suffit de disposer d'un élément χ de $\mathfrak{a}^\star$ et, si l'on note K_χ le stabilisateur de χ pour l'action (coadjointe) de K sur $\mathfrak{p}^\star$, d'une représentation irréductible μ de K_χ.

La donnée de χ et de μ permet alors de construire un sous-groupe parabolique cuspidal P_χ de G : si T est un tore maximal de K et si $A = \exp_G(\mathfrak{a})$, on peut envisager le centralisateur A_χ dans A de l'algèbre de Lie de $T \cap K_\chi$; on peut ensuite envisager le centralisateur L_χ de A_χ lui-même dans G tout entier. Il s'écrit $L_\chi = M_\chi A_\chi$, où M_χ est un sous-groupe réductif de G qui n'est compact que si la dimension de l'orbite de χ dans $\mathfrak{p}^\star$ est maximale, et qui n'est automatiquement connexe que si G est complexe, mais qui admet toujours une série discrète et dont K_χ est toujours un sous-groupe compact maximal. Un choix de système de racines positives pour la paire $(\mathfrak{g}, \mathfrak{a}_\chi)$ permet de définir un sous-groupe groupe nilpotent N_χ de G (voir [34], page 135, et bien sûr le chapitre 7), et on pose $P_\chi = M_\chi A_\chi N_\chi$.

Un théorème profond de David Vogan permet d'associer à μ une représentation tempérée et irréductible $\mathbf{V}(\mu)$ de M_χ, qui a la propriété d'être de caractère infinitésimal réel (elle peut être de série discrète, limite de série discrète, et parfois ni l'un ni l'autre). En notant ρ la demi-somme des racines positives associée au choix qui a permis de définir N_χ, et en voyant χ comme une forme linéaire sur $\mathfrak{a}_\chi^\star$, on peut obtenir une représentation tempérée de G, la représentation induite

$$\mathrm{Ind}_{M_\chi A_\chi N_\chi}^{G} \left(\mathbf{V}(\mu) \otimes e^{i\chi+\rho} \right).$$

L'un des résultats principaux du chapitre 7, et de cette thèse entière, est le suivant.

Cela produit une bijection entre le dual tempéré d'un groupe de Lie semi-simple réel et le dual unitaire de son groupe de déplacements de Cartan.

Pour le démontrer, il suffit d'utiliser les critères d'irréductibilité des représentations de G et les conditions pour que des représentations de G et de G_0 soient équivalentes, ainsi que la description du dual tempéré de G due à Knapp et Zuckerman.

Il est immédiat d'étendre cette bijection au cas des groupes de Lie linéaires connexes réductifs : je le dirai au début du chapitre 8.

3.3.b Déformation vers la courbure nulle et contraction des espaces de vecteurs lisses et des opérateurs

Dans la suite du chapitre 7, je me concentre sur les *réalisations* des représentations irréductibles (et tempérées) de G, c'est-à-dire sur les espaces (de Hilbert) où ces représentations agissent : l'une des raisons pour lesquelles les idées de Mackey ont pu être négligées depuis 1975, alors que le sous-ensemble de ces idées qui concerne les représentations de série principale (pour les valeurs régulières du caractère infinitésimal) a reçu un examen détaillé, est le peu d'intérêt de comparer les représentations de G_0 associées aux orbites de petite dimension dans $\mathfrak{p}^\star$, notamment celles qui sont associées à l'orbite zéro et sont de dimension finie, avec les représentations de dimension infinie et très difficiles à construire qui étaient alors l'objet de toutes les attentions.

Pourtant, pour définir la bijection du paragraphe précédent, je me suis servi du théorème de Vogan qui étend l'observation (due à Blattner, Harish-Chandra, Schmid) selon laquelle une représentation de série discrète a un *unique* K-type minimal. Peut-on comprendre cette relation géométriquement à l'aide de la notion de contraction (comme chez Inonü et Wigner) et des idées de Mackey ?

Pour le faire, j'introduis une famille $(G_t)_{t\in\mathbb{R}}$ de groupes de Lie ; chaque membre de la famille est isomorphe à G, sauf celui qui correspond à $t = 0$, pour lequel c'est le groupe de déplacements de Cartan G_0 ; la famille est continue en un sens précisé à la section 2 du chapitre 7, et il y a un isomorphisme explicite φ_t de G_t vers G.

Soit π une représentation irréductible et tempérée de G et $\mathcal{H}$ l'espace de Hilbert où elle agit ; si π_t est une représentation irréductible et tempérée de G_t d'espace $\mathcal{H}_t$, et si π_t et $\pi \circ \varphi_t$ sont équivalentes, alors il y a une isométrie $\mathbf{C}_t$, dont le lemme de Schur dit qu'elle est unique à un scalaire de module 1 près, de $\mathcal{H}$ vers $\mathcal{H}_t$. Dans tous les cas que j'envisagerai, j'imposerai une contrainte qui détermine $\mathbf{C}_t$ complètement.

Lorsque v est un vecteur lisse de $\mathcal{H}$, j'examine alors le comportement de $\mathbf{C}_t v$ quand t tend vers zéro. Bien sûr, pour pouvoir le faire, il faut plonger tous les $\mathcal{H}_t$ dans un espace commun.

⋆

Lorsque π est une représentation de série principale, on peut choisir pour $\mathcal{H}$ l'espace $\mathbf{L}^2(K, V)$ des fonctions de carré intégrable sur K à valeurs dans un espace V de dimension finie. C'est aussi le cas pour π_t, et alors $\mathcal{H}_t$ coïncide naturellement avec $\mathcal{H}$. D'ailleurs, c'est aussi le cas pour la représentation π_0 de G_0 que la bijection ci-dessus associe à π. Je démontre alors que pour tout v, $\mathbf{C}_t v$ a une limite quand t tend vers zéro, et que les opérateurs de π_t convergent faiblement vers ceux de π_0. Ces résultats sont très proches de ceux de Dooley et Rice, et donnent une interprétation naturelle aux changements d'échelles qui sont nécessaires sur les caractères infinitésimaux dans [18].

Compte tenu de l'importance qu'avait prise la réalisation des représentations de série principale come espaces de fonctions propres du laplacien sur G/K dans la première partie de ma thèse, j'examine ce que le résulat précédent a pour analogue lorsqu'on choisit $\mathcal{H}$ cet espace de fonctions sur G/K. Ce qui arrive alors est important pour la suite : on ne peut plus supposer que $\mathcal{H}_t$ coïncide avec $\mathcal{H}$; pour comparer v et $\mathbf{C}_t v$, il devient utile d'utiliser le fait que G_t/K est difféomorphe à $\mathfrak{p}$ pour tout t et de plonger l'espace des vecteurs lisses de chaque $\mathcal{H}_t$ dans l'espace des fonctions continues sur $\mathfrak{p}$.

J'observe alors le fait très simple suivant :

Le processus de contraction fait converger les ondes d'Helgason
vers les ondes planes ordinaires,
au sens de la topologie de la convergence uniforme sur les compacts de $\mathfrak{p}$.

À la limite quand t tend vers zéro, le processus fournit l'espace des fonctions dont la transformée de Fourier ordinaire est concentrée sur une orbite de K dans $\mathfrak{p}^\star$, qui est une réalisation pour π_0.

Ainsi, $\mathbf{C}_t v$ a une limite pour tout vecteur lisse v, mais cette fois la convergence se réfère à la topologie d'un espace de Fréchet dans lequel il est possible de plonger tous les $\mathcal{H}_t$, plutôt qu'à une topologie associée aux structures hilbertiennes sur les $\mathcal{H}_t$. Je montrerai qu'un tel espace de Fréchet peut être introduit pour (presque) *toute* représentation irréductible tempérée de G. J'expliquerai le "presque", bien sûr.

$\star$

Dans le cas plus nouveau où π est une représentation *de série discrète*, je m'appuie sur la réalisation de Parthasarathy et Atiyah-Schmid, où $\mathcal{H}$ est l'espace des sections d'un fibré vectoriel équivariant sur G/K qui sont (de carré intégrable, lisses et) *solutions de l'équation de Dirac*. Comme un tel fibré est topologiquement trivial et que G_t/K est toujours difféomorphe à $\mathfrak{p}$, je réalise alors chaque $\mathcal{H}_t$ comme un espace de fonctions (lisses) à valeurs vectorielles sur $\mathfrak{p}$ (mais l'action de G_t, issue de la trivialisation, change avec t).

Dans ce contexte, je montre que $v \mapsto \mathbf{C}_t v$ n'est autre que l'opérateur qui "zoome" sur le comportement de la fonction au voisinage de l'origine. Passer à la limite quand t tend vers zéro ne laisse alors disponible, au sens de la topologie de la convergence uniforme sur les compacts, que la valeur de la fonction en zéro.

Or, le travail d'Atiyah et Schmid interprété dans ce contexte montre que l'ensemble des valeurs en zéro possibles, qui est naturellement un K-module, est irréductible, et que sa classe d'équivalence n'est autre que le K-type minimal de π. Cela donne une explication géométrique, à l'aide de la déformation vers la courbure nulle, à l'existence du K-type minimal et à la bijection de Mackey.

Revenant à la réalisation plus habituelle comme espace de sections d'un fibré équivariant sur G/K, une explication géométrique semble donc être fournie par l'observation suivante, qui est à ma connaissance nouvelle.

Lorsqu'on réalise une représentation de série discrète comme espace des solutions d'une équation de Dirac pour les sections d'un fibré vectoriel homogène sur G/K, le sous-espace qui porte le K-type minimal est formé des sections qui sont entièrement déterminées par leur valeur au point-base $1_G K$.

$\star$

Les cas de la série principale et de la série discrète permettent d'accéder à tout le dual tempéré :

• les limites de série discrète sont construites, suivant Knapp et Zuckerman, à l'aide de produits tensoriels de représentations de série discrète et de représentations irréductibles (non-hermitiennes) de dimension finie de l'algèbre de Lie complexe ; ces dernières ont des descriptions bien connues à l'aide des modules de Verma, ce qui me permet d'utiliser les résultats sur la série discrète pour étudier les représentations limites de série discrète. C'est l'objet de la section 6 du chapitre 7.

• toutes les représentations tempérées sont induites à partir d'un sous-groupe parabolique cuspidal MAN, et de représentations de série discrète ou limites de série discrète de M, d'une façon tout à fait analogue à la construction de la série principale. On peut donc utiliser les résultats sur la série discrète et les idées et lemme rencontrés pour la série principale afin d'accéder au dual tempéré tout entier. Ce n'est pas dépourvu de technicité, parce que le groupe réductif M n'est pas connexe, et qu'il faut donc étendre les résultats sur la série discrète à la série discrète de groupes réductifs non connexes, mais ces complications techniques n'introduisent pas d'obstruction à notre stratégie et la section 7.1 du chapitre 7 s'en occupe.

Pour toute représentation tempérée irréductible π de G, on obtient donc un espace de Fréchet dans lequel plonger chacun des $\mathcal{H}_t$ ci-dessus, et pour tout vecteur lisse v de $\mathcal{H}$, il y a alors une limite (dans cet espace de Fréchet) à $\mathbf{C}_t v$ quand t tend vers zéro. L'ensemble des limites ainsi obtenues est naturellement un G_0-module, et dans presque tous les cas, c'est un G_0-module unitaire irréductible dont la classe d'équivalence est celle prédite par la bijection décrite au début du chapitre.

Les situations où notre approche donne des résultats incomplets sont celles où, dans les notations de la section 3.3.1, la représentation $\mathbf{V}(\mu)$ n'est ni de série discrète, ni limite de série discrète, ni triviale (ce dernier cas se présente exactement une fois, pour la représentation de série principale de paramètre continu nul). La raison est que je ne connais pas de réalisation géométrique adaptée de ces représentations, et que le G_0-module obtenu faute d'une réalisation géométrique spécifique n'est pas irréductible, bien qu'il contienne de façon canonique le module irréductible cherché. Cette situation ne se présente jamais pour $SL_2(\mathbb{R})$ ou pour les groupes de rang réel 1, et lorsqu'elle se présente l'ensemble des représentations concernées est de mesure de Plancherel nulle, mais cela n'enlève rien au fait que le dual tempéré n'est pas encore tout entier accessible à notre approche.

3.3.c Chapitre 8 : nouvelle démonstration de la "conjecture" de Connes-Kasparov pour les groupes de Lie réels réductifs

J'ai signalé plus haut que c'est en partant de la géométrie non-commutative que Nigel Higson avait donné une nouvelle actualité aux idées de Mackey dans le cas particulier des groupes semi-simples complexes.

Higson a en fait donné une preuve très naturelle de la *conjecture de Connes-Kasparov*[51], qui prédit la K-théorie de la $C^\star$-algèbre réduite $C^\star_r(G)$, à partir de l'analogie de Mackey et des ressemblances qu'elle induit entre $C^\star_r(G)$ et la $C^\star$-algèbre réduite $C^\star_r(G_0)$.

Dans sa démonstration, la notion de *K-type minimal* joue un grand rôle. Les K-types minimaux sont uniques pour les représentations irréductibles tempérées de groupes semi-simples complexes, et le sont aussi pour celles de leurs groupes de déplacements de Cartan ; Higson signale que bien que sa bijection ne soit pas un homéomorphisme entre les duals réduits complets, elle induit des homéomorphismes entre les sous-ensembles de $\widehat{G}_0$ et $\widehat{G}$ rassemblant les représentations de même K-type minimal.

Higson introduit alors, pour chaque élément λ de $\widehat{K}$, un sous-quotient de $C^\star_r(G)$ dont le dual est le sous-ensemble $\widehat{G}^\lambda$ des représentations tempérées de K-type minimal λ, et un sous-quotient de $C^\star_r(G_0)$ dont le dual est $\widehat{G_0}^\lambda$. L'isomorphisme de Connes-Kasparov apparaît alors comme une conséquence naturelle du fait que $\widehat{G}^\lambda$ et $\widehat{G_0}^\lambda$ sont homéomorphes, et que les sous-quotients sont Morita-équivalents à une même $C^\star$-algèbre commutative d'une façon compatible avec la déformation naturelle de G à G_0.

Ce chapitre, rédigé en décembre 2015 (arXiv :1602.08891), étend l'analyse menée par Higson aux groupes réductifs réels (supposés linéaires et connexes).

51. C'est ainsi que s'appelle la conjecture de Baum-Connes pour les groupes de Lie.

Dans le cas réel, les K-types minimaux ne sont pas uniques en général, bien qu'ils le soient pour les représentations de caractère infinitésimal réel − c'était la clé de la définition de la bijection du chapitre 7. Je commence donc par montrer le fait suivant.

La bijection du chapitre 7 est compatible avec les K-types minimaux.

Dans la suite du chapitre, je montre que cette bijection est assez naturelle pour que la démonstration d'Higson s'étende aux groupes réductifs réels sans qu'aucun changement conceptuel soit nécessaire : pour chaque sous-ensemble fini $\mathcal{C}$ de $\widehat{K}$ rassemblant les K-types minimaux d'une représentation tempérée irréductible de G, je définis un sous-quotient de $C_r^\star(G)$ dont le dual est le sous-ensemble $\widehat{G}^{\mathcal{C}}$ des représentations tempérées dont le K-type minimal est dans $\mathcal{C}$, et un sous-quotient de $C_r^\star(G_0)$ dont le dual est $\widehat{G_0}^{\mathcal{C}}$. Je montre alors que la bijection du chapitre 7 induit un homéomorphisme entre $\widehat{G}^{\mathcal{C}}$ et $\widehat{G_0}^{\mathcal{C}}$, et que les sous-quotients que je viens d'évoquer sont, ici encore, Morita-équivalents à une même $C^\star$-algèbre commutative d'une façon compatible avec la déformation de G à G_0.

La bijection décrite au chapitre 7
permet d'étendre la démonstration d'Higson de la conjecture de Connes-Kasparov
au cas des groupes de Lie réels (linéaires connexes) réductifs.

⋆ ⋆ ⋆

3.4. Un dernier mot pour ceux de mes lectrices et lecteurs qui regretteraient que ce qui suit soit en anglais, puisque nous avons, en France, le rare privilège de pouvoir encore écrire des mathématiques dans notre langue maternelle. C'est à l'impatience de voir ma thèse soutenue, aux belles occasions qui m'ont permis d'en discuter avec des collègues étrangers et d'en publier une partie déjà, que vous devez d'être en train de lire la dernière phrase en français de ce manuscrit.

Bibliographie

[1] A. Afgoustidis, *Monochromaticity of orientation maps in V1 implies minimum variance for hypercolumn size*, The Journal of Mathematical Neuroscience **5** :10 (2015)

[2] A. Afgoustidis, *Orientation maps in V1 and non-Euclidean geometry*, The Journal of Mathematical Neuroscience **5** :12 (2015)

[3] D. Angelaki and K. Cullen, *Vestibular system : the many facets of a multimodal sense*,Annu Rev Neurosci. **31** (2008), pp. 125-150.

[4] M. Atiyah, G.L. Luke (Editors), *Representation Theory of Lie Groups*, Proceeedings of the SRC/LMS Research Symposium on Representations of Lie Groups, Oxford, June-July 1977. London Mathematical Society Lecture Note Series **34**, Cambridge University Press, 1979.

[5] V. Bargmann, *Irreducible unitary representations of the Lorentz group*, Annals of Mathematics **48(3)** (1947), pp. 568-640.

[6] V. Bargmann, *On unitary ray representations of continuous groups*, Annals of Mathematics **59(1)** (1954), pp. 1-46.

[7] D. Bennequin, *Remarks on invariance in the primary visual system of mammals*, in ; Neuromathematics of Vision, A. Sarti and G. Citti (eds), pp. 243-333. Springer Lecture Notes in Morphogenesis Series (2014).

[8] D. Bennequin, R. Fuchs, A. Berthoz and T. Flash, *Movement timing and invariance arise from several geometries*, PLoS Computational Biology **5(7)** (2009), pp. 1-27.

[9] D. Bennequin and A. Berthoz, *Non-Linear Galilean vestibular receptive fields*, Conf Proc IEEE Eng Med Biol Soc (2011), pp.2273-2276.

[10] A. Berthoz, *Le sens du mouvement*, Odile Jacob (1997).

[11] W. Bosking, W. Zhang, B. Schofield and D. Fitzpatrick, *Orientation selectivity and the arrangement of horizontal connections in tree shrew striate cortex*, The Journal of neuroscience **17(6)** (1997), pp. 2112-2127.

[12] E. Cassirer, *Le concept de groupe et la théorie de la perception*, Journal de Psychologie normale et pathologique **35** (1938), pp. 368-414.

[13] J.-L. Clerc, *Une formule asymptotique du type Mehler-Heine pour ¡es zonales d'un espace riemannien symétrique*, Studia Math. **57** (1976), pp. 27-32

[14] R. David, A. Berthoz and D. Bennequin, *Secret laws of the labyrinth*, Conf Proc IEEE Eng Med Biol Soc. (2011), pp. 2269-72

[15] P. Dayan & L. Abbott, *Theoretical Neuroscience : Computational and Mathematical Modeling of Neural Systems.* MIT Press, 2001.

[16] M. Dimiccoli, B. Girard, A. Berthoz and D. Bennequin. *Striola magica. A functional explanation of otholith organs*, Journal of Computational Neuroscience **35(2)** (2013), pp. 125-54

[17] P. Dirac, *Relativistic wave equations*, Proc. of the Royal Society of London. Series A, Vol. **155**, No. 886 (1936), pp. 447-459.

[18] A. Dooley and J. Rice, *On Contractions of Semisimple Lie Groups*, Transactions of the American Mathematical Society **289(1)** (1985), pp. 185-202

[19] C. Eckes, *Les groupes de Lie dans l'oeuvre de Hermann Weyl, traduction et commentaires de la « Théorie de la représentation des groupes continus semi-simples par des transformations linéaires » (1925-1926)*, Presses universitaires de Nancy -Editions universitaires de Lorraine, coll. "Histoires de géométries" (2014).

[20] P. Etingof (dir.), *Introduction to representation theory (with historical interludes by Slava Gergovitch)*, AMS Student Mathematical Library (2011).

[21] S. Gelbart, *An elementary Introduction to the Langlands Program*, Bull. Amer. Math. Soc. **10** (1984), pp. 177-219

[22] H. Gray, illustrated by H. Carter, *Anatomy of the Human Body*, Churchill Livingstone (1918).

[23] T. Hawkins, *Emergence of the theory of Lie groups*, Sources and Studies in the history of Mathematics and the Physical Sciences, Springer (2000).

[24] S. Helgason, *Geometric analysis on symmetric spaces.* American Mathematical Society ; 2008.

[25] H. von Helmholz, *Uber die Tatsachan, die der Geometrie zu Grunde liegen*, Nachrichten von der königlichen Gesellschaft der Wissenschaften zu Göttingen **9** (1868). Trad. in *Epistemological writings (ed. P. Hertz and M. Schlik, trad. M. Lowe), Springer, 1977.*

[26] R. Herb, *Harish Chandra and his work*, Bull. Amer. Math. Soc **25(1)** (1991), pp. 1-18.

[27] R. Howe, *Harish-Chandra, 1923-1983*, Biographical Memoirs, National Academy of Sciences

[28] D. Hubel and T. Wiesel, *Ferrier lecture. Functional architecture of macaque monkey visual cortex*, Proc R Soc Lond B Biol Sci. **f198(1130)** (1977), pp. 1-59

[29] D. Hubel, Evolution of Ideas on the Primary Visual Cortex, 1955-1978 : a Biased Bistorical Account, Nobel Lecture (December 8, 1981).

[30] E. Inonu, *A historical note on group contractions*, 1997.

[31] E. Kandel, J. Schwartz, and T. Jessell (Editors). *Principles of neural science.* New York, McGraw-Hill ; 2000.

[32] E. Kaniuth and K. Taylor, *Induced representations of locally compact groups.* Cambridge Tracts in Mathematics **197**, Cambridge University Press (2013).

[33] M. Kaschube, M. Schnabel, S. Lowel, D. Coppola, L. White and F. Wolf, Online Supplementary Material of : *Universality in the evolution of orientation columns in the visual cortex*, Science, **330(6007)** (2010), pp. 1113-1116.

[34] A. W. Knapp, *Representation theory of semisimple groups : An overview based on examples.* Second edition. Princeton Landmarks in Mathematics **36**, Princeton University Press (2001).

[35] R. P. Langlands, *Harish-Chandra Memorial Talk*, in The mathematical legacy of Harish-Chandra (edited by R. Doran and V. S. Varadarajan), American Mathematical Society (2000), pp. 47-49.

[36] R. P. Langlands, *Harish-Chandra (1923-1983)*, Biog. Mem. of Fellows of the Roy. Soc. **31** (1985).

[37] J. M. Lévy-Leblond, *Une nouvelle limite non-relativiste du groupe de Poincaré.* Annales de l'institut Henri Poincaré (A) Physique théorique **3(1)** (1965) : 1-12

[38] M. Lévy-Nahas, *Deformation and Contraction of Lie Algebras*, Journal of Mathematical Physics **8(6)** (1967),1211-1222

[39] G. W. Mackey, *A theorem of Stone and von Neumann.* Duke Math. Journal **16(2)** (1949), pp. 313-326

[40] G. W. Mackey, *Imprimitivity for Representations of Locally Compact Groups I*, Proc Natl Acad Sci U S A. **35(9)** (1949), pp. 537–545.

[41] G. W. Mackey, *Theory of Unitary Group Representations*, Chicago Lectures in Mathematics, Chicago University Press, 1976.

[42] G. W. Mackey, *Harmonic analysis as the exploitation of symmetry-a historical survey*, Bulletin of the American Mathematical Society **3(1)** (1980), 543-698.

[43] J. Mickelsson and J. Niederle, Contractions of representations of de Sitter groups, Communications in Mathematical Physics **27(3)** (1972), pp 167-180.

[44] N. Mukunda, *Expansion of Lie Groups and Representations of SL(3, C)*, J. Math. Phys. 10, 897 (1969)

[45] K. Okhi, S. Chung, P. Kara, M. Hubener, T. Bonhoeffer and R. Reid, *Highly ordered arrangement of single neurons in orientation pinwheels.* Nature **442(7105)** (2006), pp. 925-928.

[46] J. Petitot, *Neurogéométrie de la vision : modeles mathematiques et physiques des architectures fonctionnelles.* Paris : Editions Ecole Polytechnique ; 2008.

[47] Q. C. Pham and D. Bennequin, *Affine invariance of human hand movements : a direct test*, arXiv preprint (2012) : arXiv :1209.1467.

[48] J. Piaget et B. Inhelder, *La représentation de l'espace chez l'enfant*, Presses universitaires de France, 1947.

[49] H. Poincaré, *La science et l'hypothèse*, Paris, 1902.

[50] L. Pontryagin, *Sur les groupes abéliens continus*, Comptes-rendus de l'académie des Sciences **198** (1934), pp. 328-330

[51] J. Ribot, A. Romagnoli, C. Milleret, D. Bennequin and J. Touboul, *Pinwheel-dipole configuration in cat early visual cortex*, Neuroimage **128** (2015), pp. 63-73.

[52] A. Romagnoli, J. Ribot, D. Bennequin and J. Touboul, *Parsimony, Exhaustivity and Balanced Detection in Neocortex*, PloS Computational Biology **11(11)**, 2015.

[53] F. Rouvière, *Symmetric Spaces and the Kashiwara-Vergne Method*, Lecture Notes in Mathematics **2115** (2014), Springer Berlin Heidelberg.

[54] E. Saletan, *Contraction of Lie groups*, Journal of Mathematical Physics **2(1)** (1961), pp. 1-21

[55] E. Schrödinger, *Quantisation as a Problem of Proper Values*, Annalen der Physik **79** (1926), pp. 1-12.

[56] D. Shale, *Linear symmetries of free boson fields*, Trans. Amer. Math. Soc **103** (1962), pp. 149-167.

[57] I. Segal, *A class of operator algebras which are determined by groups*, Duke Math. Journal **18(1)** (1951), 221-265.

[58] J-M. Souriau, *Grammaire de la Nature*, 2007.

[59] S. Sternberg, *Group theory and physics.* Cambridge University Press, 1995.

[60] V. S. Varadarajan, *Geometry of Quantum Theory*, second edition. Springer Science & Business Media, 2007.

[61] V. S. Varadarajan, *An Introduction to Harmonic Analysis on Semisimple Lie Groups.* Cambridge studies in advanced mathematics **16**, Cambridge University Press, 1989.

[62] N. Ja. Vilenkin and A. U. Klimyk, *Representations of Lie groups and special functions.* Three volumes. Kluwer Academic Publishers, 1993.

[63] J. von Neumann, *Almost periodic functions in a group*, Trans. Amer. Math. Soc. **36** (1934), pp. 445-492.

[64] A. Weil, *L'intégration dans les groupes topologiques et ses applications à l'analyse*, Hermann, 1940.

[65] A. Weil, *Sur certains groupes d'opérateurs unitaires*, Acta Mathematica **111** (1964), pp. 143-211.

[66] H. Weyl, *The theory of groups and quantum mechanics*, 1929. English Translation : Dover, 1950.

[67] H. Weyl, *Gruppentheorie und quantenmechanik*, Zeitschrift fur Physik **46(1)** (1927), pp. 1-46.

[68] E. Wigner, *On unitary representations of the inhomogeneous Lorentz group*, Annals of Mathematics **40(1)** (1939), pp. 149–204

[69] M. Schnabel, M. Kaschube, S. Löwel and F. Wolf, *Random waves in the brain : Symmetries and defect generation in the visual cortex*, Eur. Phys. J. Special Topics **145** (2007), pp. 137-157.

Part I

Some questions regarding harmonic analysis and random fields on homogeneous spaces, motivated by the abundance of symmetry arguments in the study of the primary visual cortex.

Chapter 1

A Moiré pattern on symmetric spaces of noncompact type

Contents

Abstract. I prove that if X is a symmetric space of the noncompact type, just as adding Helgason waves which propagate in all direction yields an elementary spherical function for X, a Helgason wave can be produced by adding elementary spherical functions whose centers describe a horocycle in X.

1 Introduction

A *moiré pattern* is a visual effect obtained by superimposing plane motifs which are obtained from one another through small Euclidean motions. Moiré patterns often occur in image processing: see [7, 1], but they also appear in other contexts. Let me start by describing a possible use in Neuroscience [11, 12, 13] which is the motivation for this chapter.

I recalled in the Introduction that on the way from the retina to the primary visual cortex, the visual information is conveyed by the Lateral Geniculate Nucleus (hereafter abridged as LGN). The response of a lateral geniculate cell to the visual input can be described with the help of a receptive profile − a function R_{LGN} defined on the visual plane, as in section 1.2(a) above.

It is well-acknowledged that the receptive profiles of LGN cells have *spherical* symmetry: to each cell is attached a point x_0 of the visual plane, and the receptive profile R_{LGN} is a function of the distance to x_0. In addition, as the distance to x_0 grows, R_{LGN} decreases to zero, then becomes negative in a region in which the presence of light has an inhibitory effect on the given cell, then grows again, and tends to zero as the distance to x_0 grows again. A famous suggestion for R_{LGN} is a mexican hat function, that is, the Laplacian of a Gaussian function.

In the primary visual cortex, however, the receptive profiles do not have spherical symmetry: as I recalled in the Introduction, they have a preferred direction, and the natural candidates for the receptive profiles are products of a plane wave with a function which decreases with the distance to a preferred position (popular models include Gabor wavelets discussed in the Introduction, or the product of a plane wave with a mexican hat function, known as a Marr wavelet).

The transition from the LGN to V1 then involves a change in the symmetry of the receptive profiles. What is the biological basis for this transition ?

Hubel and Wiesel famously proposed that the answer lies in the wiring of neurons: if a given V1 neuron receives inputs from LGN cells which have their centers of symmetry *aligned and close to one another*, and if the combination of LGN inputs is a simple summation, then a directional preference can emerge through a Moiré-like pattern .

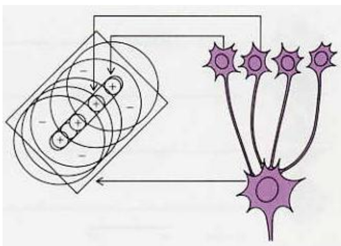


Figure 1: Hubel and Wiesel's scenario for the transition between the receptive profiles in the LGN and those in V1: if the electrical activity of a neuron in V1 is close to the sum of activities exhibited by LGN cells which have the centers of their receptive profiles distributed on a small segment of the visual plane, the cortical neuron will have a clear orientation preference.

It is possible to emphasize the role of symmetries in Hubel and Wiesel's argument, by changing the models for the receptive profiles and relaxing the condition that the receptive profiles decrease at infinity. In fact, if we drop that realistic but symmetry-independent requirement, a natural mathematical counterpart to Hubel and Wiesel's argument sits inside the structure of the Euclidean group. I shall now say what the symmetry-based counterpart is first, and then indicate how it is related to the structure of the Euclidean motion group.

As our translation-invariant receptive profile, choose a plane wave $x \mapsto e^{i\langle Ru, x\rangle}$, where R is a positive number and u a unit vector in $\mathbb{R}^2$. As our rotation-invariant receptive profile, choose the Bessel function $J_R := x \mapsto \int_{\mathbb{S}^1} e^{i\langle Ru,x\rangle} du$ (here the Haar measure on $\mathbb{S}^1$ is normalized so as to have total mass one).

I claim that a plane wave can be reconstructed by the constructive intereference of Bessel functions whose centers of symmetry are distributed on a straight line whose direction is orthogonal to that of the wave's propagation. Let us first observe a picture of the constructive interference (Figure 2).

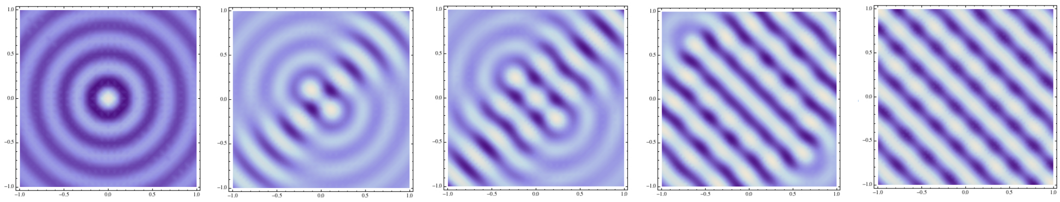


Figure 2: On the left is a Bessel function; left to right are superpositions of 3, 5, 11, 21 Bessel functions whose centers of symmetry lie along the obvious line.

Now, suppose $u^\perp$ is a unit vector orthogonal to u. I shall now indicate how the elementary properties of the Fourier transform, in particular the fact that the Fourier transform of the Dirac distribution on the line $\mathbb{R}u$ is the Dirac distribution on the orthogonal line $\mathbb{R}u^\perp$, imply

$$\int_{\mathbb{R}} J_R(x+tu)dt = \frac{2}{R}\cos\left(R\langle x, u^\perp\rangle\right), \tag{1.1}$$

an apparently natural guess in view of Figure 2.

However, the equality cannot hold in the usual, strong sense: J_R is not an integrable function (if it were its Fourier transform would be continuous, and it it the Dirac distribution on the circle of radius R), and the left-hand side of (1.1) is not an absolutely convergent integral. The left-hand-side does have a meaning as an improper integral, because when x and u are fixed, we have

$$\begin{aligned}\int_{-A}^{A} J_R(x+tu)dt &= \int_{[-A,A]} dt \int_{\mathbb{S}^1} e^{iR\langle x+tu, v\rangle} dv = \int_{\mathbb{S}^1} e^{iR\langle x,v\rangle} \int_{[-A,A]} dt e^{iR\langle tu,v\rangle} \\ &= 2\int_{\mathbb{S}^1} e^{iR\langle x,v\rangle} \frac{\sin\left(AR\langle u,v\rangle\right)}{R\langle u,v\rangle} dv,\end{aligned}$$

and this actually has a limit as A goes to infinity. But because of the stationary-phase lemma (applicable here because $v \mapsto \langle u, v\rangle$ admits only two critical points on $\mathbb{S}^1$, and that these are nondegenerate), there is a constant ℓ such that $\int_{-A}^{A} J_R(x+tu)dt \underset{A\to\infty}{\sim} \frac{\ell}{\sqrt{A}}$, so the limit is zero!

If we are to interpret Figure 2 with the help of (1.1), we have to find a weaker meaning for the left-hand side. I shall now argue that it is best to interpret (1.1) as an equality of distributions.

When ψ is a Schwartz function on $\mathbb{R}^2$, we can write

$$\int_{x+\mathbb{R}u} \psi J_R = \int_{(x+\mathbb{R}u)\times\mathbb{S}^1} \vartheta(y) e^{Ri\langle y,v\rangle} dy dv = \int_{\mathbb{S}^1} \mathcal{F}\left(\psi \delta_{x+\mathbb{R}u}\right)[Rv] dv, \tag{1.2}$$

where δ_{x+Ru} is the Dirac distribution on the line $x + \mathbb{R}u$ and $\mathcal{F}$ is the Euclidean Fourier tranfsorm. I shall assume the position x and the direction u to be fixed here.

Now choose a family $(\vartheta_\varepsilon)_{\varepsilon>0}$ of Schwartz functions on $\mathbb{R}^2$ which, as ε goes to zero, goes in the space $\mathcal{S}'(\mathbb{R}^2)$ of tempered distributions on $\mathbb{R}^2$ to the Dirac distribution $\delta_{x+\mathbb{R}u}$ over the line $D = x + \mathbb{R}u$: one can for instance start from a smooth, nonnegative-valued, compactly supported function ϖ on $\mathbb{R}$ which is identically one in a neighborhood of the identity and has integral one, and then set $\vartheta_\varepsilon(y) = \frac{1}{\varepsilon}\varpi\left(\frac{d(y,D)}{\varepsilon}\right)\varpi\left(\varepsilon d(y,0)\right)$, where d is the Euclidean distance in $\mathbb{R}^2$.

Then as ε goes to zero, the Fourier transform $\mathcal{F}\left(\vartheta_\varepsilon\right)$ goes to that of $\delta_{x+\mathbb{R}u}$, which is the product $\left(\xi \mapsto e^{i\langle \xi,x\rangle}\right)\delta_{\mathbb{R}u^\perp}$ between a plane wave and the Dirac distribution on $\mathbb{R}u^\perp$.

Assigning to a tempered distribution T on $\mathbb{R}^2$ the distribution on $\mathbb{R}_\star^+$ which sends a smooth and compactly supported function α on $\mathbb{R}_\star^+$ to the number $\langle T, \tilde{\alpha}\rangle$, where $\tilde{\alpha}$ is the radial function on $\mathbb{R}^2$ built on α and the bracket is the duality bracket, we obtain a map I from $\mathcal{S}'(\mathbb{R}^2)$ to the space $\mathcal{D}'(\mathbb{R}_\star^+)$ of distributions on $\mathbb{R}_\star^+$. Noticing by a polar change of coordinates that $R \mapsto R\left(\int_{\mathbb{S}^1} \mathcal{F}\left(\vartheta_\varepsilon\right)[Rv]dv\right)$ is the image under I of $\mathcal{F}(\vartheta_\varepsilon)$, and that I is continuous with respect to the natural topologies of $\mathcal{S}'(\mathbb{R}^2)$ and $\mathcal{D}'(\mathbb{R}_\star^+)$, we thus see that $R \mapsto \int_{\mathbb{S}^1} \mathcal{F}\left(\vartheta_\varepsilon\right)[Rv]dv$ has a limit in $\mathcal{D}'(\mathbb{R}_\star^+)$ as ε goes to zero. The previous calculation shows that the limit is in fact the continuous map

$$R \mapsto \frac{1}{R}\int_{\mathbb{S}^1} e^{i\langle x,v\rangle}\delta_{\mathbb{R}u^\perp}[Rv]dv = \frac{1}{R}e^{i\langle x,Ru^\perp\rangle} + \frac{1}{R}e^{i\langle x,-Ru^\perp\rangle} = \frac{2}{R}\cos\left(R\langle x,u^\perp\rangle\right).$$

We can thus interpret (1.1) as identifying the limit, in $\mathcal{D}'(\mathbb{R}_\star^+)$, of $R \mapsto \int_{\mathbb{R}^2} \vartheta_\varepsilon J_R$ as ε goes to zero in $\mathbb{R}$ and thus ϑ_ε goes to $\delta_{x+\mathbb{R}u}$ in $\mathcal{S}'(\mathbb{R}^2)$. This may seem far-fetched, but Figure 2 is there to remind us that the interpretation might be rather convincing; in addition, the fact that the biological receptive profiles R_{LGN} do rapidly decrease at infinity makes it all the more natural in our context to consider (1.2) before going over to (1.1).

⋆

I said above that the plane waves and the Bessel function J_R sit inside the structure of the Euclidean motion group: if one starts with the space of smooth (and, say, bounded) functions on $\mathbb{R}^2$, equipped with the natural action of the Euclidean motion group, and if one looks for the invariant subspaces, the space of functions whose Fourier transform is concentrated on a circle of radius R appears as an irreducible invariant subspace (for more details, see Chapter 3, section 2.4). The above special functions are the only elements in that space which are invariant under a one-dimensional Lie subgroup of the Euclidean group: the plane wave $x \mapsto e^{i\langle Ru,x\rangle}$ is, along with its conjugate and the linear combinations of the two, the only element invariant under $\mathbb{R}u$, and and the Bessel function $x \mapsto J_R(x - x_0)$ is (up to a scalar multiplication) the only function invariant under the subgroup of rotations around x_0.

In this short chapter, I show that the tools of non-commutative harmonic analysis make it possible to exhibit a similar Moiré pattern on a special class of negatively-curved homogeneous spaces − the symmetric spaces of noncompact type.

Soon after I had submitted my manuscript, François Rouvière pointed out to me that the initial version of this chapter was seriously flawed: the integral in the left-hand-side of (5.1) is, at least for $\lambda = 0$, a divergent one. Only then did I realize that both (1.1) and (5.1) have to be given a weaker meaning. In addition, my proof of the Lemma in section 5 was incorrect, the initial statement of (5.1) had the wrong kind of W-invariance, and though I had of course been unaware of its presence in earlier work, equation (5.5) below appears as eq. (70) at the bottom of page 219 of the 2008 edition of [10]. I thank François Rouvière for pointing this out to me, as well as providing me with a self-contained proof of the Lemma and several detailed comments which led to a major revision of this chapter.

2 Notations

In this chapter, G will be a real, connected, noncompact semisimple Lie group with finite center. Let me introduce some usual notations :

– I will write $\mathfrak{g}$ for the Lie algebra of G, fix a maximal compact subgroup K, and choose Lie subgroups A and N (the Lie algebra of A will read $\mathfrak{a}$) so that $G = KAN$ is an Iwasawa decomposition of G. The choice of N comes with a choice of positive root system for the pair $(\mathfrak{g}, \mathfrak{a})$; I shall write ρ for the corresponding half-sum of positive roots, and $\mathfrak{a}^\star_+$ for the corresponding (open) positive Weyl chamber in the dual $\mathfrak{a}^\star$.
– I shall write M for the centralizer of A within K, and B for the compact quotient K/M.
– I will assume G-invariant measures to have been chosen on G and the various subgroups and quotients in a coherent manner (see Helgason [9]). The integrations to come will be performed with respect to these invariant measures, lest some precision be given. I assume the Haar meaure of K to be normalized in such a way that the volume of B is one.

Let X be the riemannian symmetric space (of the noncompact type) G/K. Let us follow Helgason [10] in calling the orbit in X of any conjugate of N a *horocycle* (Poincaré used to say *horisphère*). If $g = k_0 a_0 n_0$ is in G, the subgroup $gNg^{-1} = k_0 N k_0^{-1}$ depends only on the image $b_0 = k_0 M$ of k_0 in B; let me say that the orbits of gNg^{-1} in X have *direction* b_0.

When x is in X and b is in B, let me write $\xi(b, x)$ for the horocycle through x with direction b; it is the orbit of x under the N-conjugate corresponding to b as above.

Suppose b is in B and $x = naK$ is in X; let me use the Iwasawa projection $\mathcal{A} : G \mapsto \mathfrak{a}$ (defined as $nak \mapsto log_A(a)$) and set $\Delta(b, x) = \mathcal{A}(b^{-1}\tilde{x}) \in \mathfrak{a}$, where $\tilde{x}$ is any lift of x in G. The element $\Delta(b, x)$ of $\mathfrak{a}$ depends only on the horocycle $\xi(b, x)$: when $\xi(b, x)$ and $\xi(b', x')$ coincide, so do $\Delta(b, x)$ and $\Delta(b', x')$.

When G equals $SU(1, 1)$ and acts through homographies on the open unit disk $\mathbb{D}$ in $\mathbb{C}$, the stabilizer for the origin 0 is a maximal compact subgroup K of G, which is isomorphic to $SO(2)$; the horocycles in $\mathbb{D}$ (whose above definition depends only on the choice of K) are the circles which are tangent to $\mathbb{D}$ at a point of its boundary; it is then natural to identify the direction of a horocycle with the tangency point.

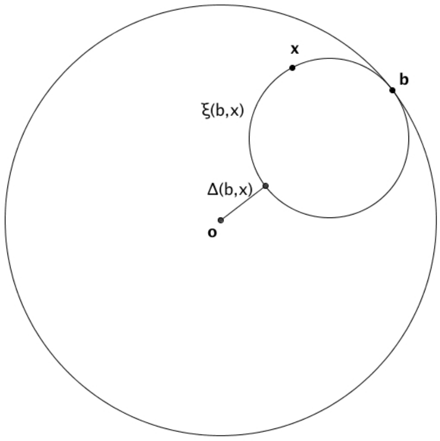


Figure 3: A horocycle in $\mathbb{D}$.

3 The Fourier-Helgason transform on X

Suppose λ is in $\mathfrak{a}^\star$ and b is in B. Set

$$\begin{aligned} e_{\lambda,b} : X \to& \mathbb{R} \\ x \mapsto& e^{\langle i\lambda+\rho \mid \Delta(b,x)\rangle}. \end{aligned}$$

The function $e_{\lambda,b}$ takes a single value on each horocycle with direction b. It is a building block for G-invariant harmonic analysis on X in much the same way as plane waves are for Fourier analysis on Euclidean space.

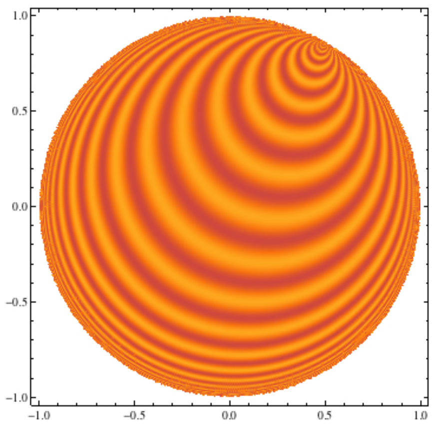


Figure 4: This is a plot of the *phase* levels of a Helgason wave (the growth factor has been deleted).

When f is a map from X to $\mathbb{C}$, the Fourier-Helgason transform of f is the map

$$\hat{f} : (\lambda \in \mathfrak{a}^\star, b \in K/M) \mapsto \int_X e_{-\lambda,b}(x) f(x) dx,$$

defined on the subset of $\mathfrak{a}^\star \times K/M$ where the above integral converges; it is for instance defined on all of $\mathfrak{a}^\star \times K/M$ if f is a smooth function with compact support. When the hat is too short for the notation to be legible, I will write $\mathcal{F}(f)$ instead of $\hat{f}$.

When f is an integrable function on X, it is no longer obvious that this integral should converge for $(\lambda, b) \in \mathfrak{a}^\star \times K/M$; yet one can show ([10], p. 209) that it does converge for (λ, b) in $\mathfrak{a}^\star \times B_0$, with B_0 a full-measure subset of B.

In what follows, I will need an extension of the Fourier-Helgason transform to distributions on X, and a non-Euclidean analogue of Schwartz-class functions and of tempered distributions. In order to get to the point more quickly, I have relegated the corresponding definitions of $\mathcal{S}(X)$, $\mathcal{S}'(X)$ and their counterparts over $\mathfrak{a} \times K/M$ to section 6 below (see [10], around p. 214). But even when f is only integrable, when $\hat{f}$ is integrable with respect to the measure $(|\mathbf{c}(\lambda)|^{-2}d\lambda) \otimes db$ on $\mathfrak{a}^\star \times K/M$ (the measure features Harish-Chandra's $\mathbf{c}$-function), the following inversion formula will hold for almost every x in X :

$$f(x) = \frac{1}{|W|} \int_{\mathfrak{a}^\star \times B} \hat{f}(\lambda, b)\ e_{\lambda,b}(x)\ |\mathbf{c}(\lambda)|^{-2} d\lambda db.$$

Here $|W|$ is the order of the Weyl group $W(\mathfrak{g}, \mathfrak{a})$; the $\mathbf{c}$-function is real-analytic on the complement in $\mathfrak{a}^\star$ of a finite union of hyperplanes, but I will not need the details of its definition. When shifting from Euclidean space to symmetric spaces of the noncompact type, there are often ways to translate interesting questions about the usual Fourier transform (like the Plancherel formula, the Paley-Wiener theorem...) into questions on the Fourier-Helgason transform, and the answers often show some likeness in spite of some important differences due to the curvature of X and the ensuing growth at infinity of Helgason's waves.

4 Elementary spherical functions

Helgason's waves $e_{\lambda,b}$ make G-invariant harmonic analysis on G/K, a subject depicted in detail in his work, look familiar; before Helgason made it look so, the fact that the function obtained by constructive interference from all Helgason waves with frequency λ is the spherical function φ_λ had already proved to be a key point in Harish-Chandra's program for studying the reduced dual of G (see [6]).

For each $\lambda \in \mathfrak{a}^\star$, the map

$$\begin{aligned} \varphi_\lambda : G &\to \mathbb{C} \\ g &\mapsto \int_B e_{\lambda,b}(gK)\ db \end{aligned}$$

takes the value 1 at zero, is left-and-right K-invariant and is and an eigenfunction for all G-invariant differential operators on G: it is an elementary spherical function of G. The only functions with the three properties in the previous sentence are the φ_λ, $\lambda \in \mathfrak{a}^\star$, and two functions of this type are equal if and only if the elements of $\mathfrak{a}^\star$ defining them are on the same orbit for the action of the Weyl group on $\mathfrak{a}^\star$.

Although I shall not need the fact itself except through several of its consequences, let me recall that the K-invariant version of Fourier-Helgason analysis led Harish-Chandra to define the $\mathbf{c}$-function and opened him the way towards the Plancherel formula for G: if f is a smooth function on G which has compact support and is K-bi-invariant, let us write $\tilde{f}(\lambda) = \int_G f(g)\varphi_{-\lambda}(g)dg$ for $\lambda \in \mathfrak{a}^\star$; then

$$f \mapsto \tilde{f} \text{ extends to an isometry between } \mathbb{L}^2(K\backslash G/K) \text{ and } \mathbb{L}^2\left(\mathfrak{a}^\star_+, |\mathbf{c}(\lambda)|^{-2}d\lambda\right).$$

5 A moiré pattern

In the next few paragraphs, I am going to prove that a synthesis formula holds in the opposite direction and that Helgason's waves can be recovered by constructive interference from spherical functions whose centers of symmetry cluster along a horocycle.

To be precise, let me choose a "frequency" λ in $\mathfrak{a}^\star$ and a point in the boundary – say the identity coset $b_0 = 1_K M$ in B. When y is in X, let me write $\varphi_\lambda^{[y]}$ for the only element in the eigenspace $\mathcal{E}_\lambda(X)$ from [10], chapter 6 [1] which takes the value 1 at y and is insensitive to left-and right-translations of the variable along an element of the stabilizer of y in G. I am going to argue that for every x in X, the equality [2]

$$\int_{\xi(b_0,0)} \varphi_\lambda^{[y]}(x)\, dy = \frac{|\mathbf{c}(\lambda)|^2}{|W|} \sum_{w \in W} e_{w\lambda, b_0}(x) \tag{5.1}$$

holds in a weak sense analogous to (1.1): here is the statement of what I am actually going to prove.

Choose a family $(\vartheta_\varepsilon)_{\varepsilon>0}$ of Schwartz functions on X which, as ε goes to zero, goes in $\mathcal{S}'(X)$ to the Dirac distribution over the horocycle $\xi(b_0, x)$ – an example is $\vartheta_\varepsilon(y) = \frac{1}{\varepsilon^{\dim(A)}} \varpi \left(\frac{d(y, \xi(b_0, x))}{\varepsilon} \right) \varpi \left(\varepsilon d(y, o)^2 \right)$, where d is the Riemannian distance in X and ϖ is the bump function introduced above.

Proposition. *When x is a point in X, the continuous function $\lambda \in \mathfrak{a}^\star \mapsto \frac{|\mathbf{c}(\lambda)|^2}{|W|} \sum_{w \in W} e_{w\lambda, b_0}(x)$ is the limit, in $\mathcal{D}'(\mathfrak{a}^\star / W)$, of $\lambda \mapsto \int_{\xi(b_0,0)} \vartheta_\varepsilon(y) \varphi_\lambda^{[y]}(x)\, dy$ as ε goes to zero.*

Note that for every $\varepsilon > 0$, $\lambda \mapsto \int_{\xi(b_0,0)} \vartheta_\varepsilon(y) \varphi_\lambda^{[y]}(x)\, dy$ is a well-defined, continuous, W-invariant function of λ.

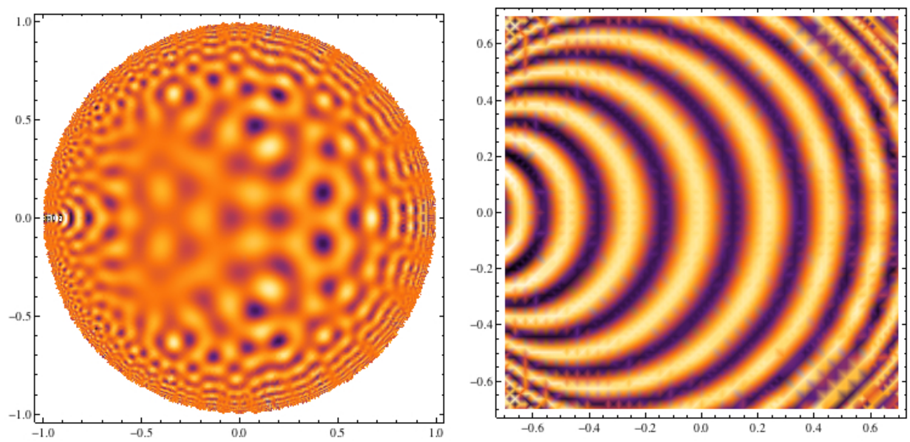


Figure 5: The left picture shows a sum of five, and the right picture the detail of a sum of sixty, spherical functions whose centers of symmetry lie on the horocycle $\xi(-1, 0)$.

1. This is the common eigenspace for G-invariant differential operators on X which contains φ_λ.
2. In the right-hand side of (5.1), the averaging over W mirrors the appearance of an even, real-valued cosine wave (rather than a complex-valued plane wave) in (1.1), and the presence of the $\mathbf{c}$-function mirrors the normalization come from the polar change of coordinates used for proving (1.1) (see section 4).

If $y \in X$ is where the origin $o = eK$ in G/K is sent by $g_y \in G$, then the function $\varphi_\lambda^{[y]}$ is none other than $z \mapsto \varphi_\lambda(g_y^{-1} \cdot z)$. When y is on the horocycle $\xi(b_0, o)$, the element g_y can be assumed to belong to N and $g_y^{-1} \cdot x$ is on the horocycle $\xi(b_0, x)$. So what I need to prove is a weak version of

$$\int_{\xi(b_0,x)} \varphi_\lambda = |\mathbf{c}(\lambda)|^2 \sum_{w \in W} e_{w\lambda, b_0}(x). \tag{5.2}$$

But disregarding the convergence questions for a few lines, the reconstruction formula by Harish Chandra cited above (see also [5]) formally yields :

$$\int_{\xi(b_0,x)} \varphi_\lambda = \int_{\xi(b_0,x)} dy \int_B e_{\lambda,b}(y) db. \tag{5.3}$$

If we swap the integrals in (5.3), we will end up with the integral over B of the Fourier-Helgason transform of the indicatrix of the horocycle $\xi(b_0, x)$, and this will bring us very close to the desired conclusion. But we will also end up with divergent integrals. Before we address this, we record the following lemma:

Lemma : *the Helgason-Fourier transform of the Dirac distribution on the horocycle $\xi(b_0, x)$ (viewed as a tempered distribution on $\mathfrak{a}^\star \times B$) is:*

$$[\lambda \mapsto e_{\lambda,b_0}(x)] \otimes \delta_{b=b_0}. \tag{5.4}$$

This is an analogue of the projective property of the Fourier transform on Euclidean space used in section 1 above, but it should be noted that the privileged direction in the Helgason-Fourier transform is that of the horocycle itself rather than an "orthogonal" one.

Proof (pointed out to me by François Rouvière): Let me write $T_{b_0,x} \in \mathcal{S}'(\mathfrak{a}^\star \times B)$ for the distribution (5.4). What I have to check is that for every ψ in $\mathcal{S}(X)$,

$$\langle T_{b_0,x} | \widehat{\psi} \rangle = \int_{\xi(b_0,x)} \psi,$$

in other words

$$\int_{\mathfrak{a}^\star} e_{\lambda,b_0}(x) \widehat{\psi}(\lambda, b_0) d\lambda = \int_{\xi(b_0,x)} \psi. \tag{5.5}$$

Let me first assume ψ to have compact support and set out from the fact that $\widehat{\psi}(\lambda, b_0) = \int_X e^{\langle -i\lambda+\rho, \mathcal{A}(y) \rangle} \psi(y) dy$. Let's use the integration formula on page 266 of [9]: if f lies in the space $\mathcal{D}(X)$ of smooth and compactly supported functions, then

$$\int_X f(y) dy = \int_{\mathfrak{a}} e^{-2\langle \rho, H \rangle} dH \int_N f(ne^H \cdot o) dn.$$

Setting $F(H) = \int_N \psi(ne^H \cdot o) dn$ for H in $\mathfrak{a}$, we obtain

$$\widehat{\psi}(\lambda, b_0) = \int_{\mathfrak{a}} e^{\langle -i\lambda+\rho, H \rangle} F(H) dH$$

(in words, the Fourier transform of ψ is the Euclidean Fourier transform on $\mathfrak{a}$ of its Radon transform over the family of horocycles with direction b). As a consequence, we obtain

$$e_{\lambda,b_0}(x) \widehat{\psi}(\lambda, b_0) = \int_{\mathfrak{a}} e^{\langle i\lambda+\rho, \mathcal{A}(x)-H \rangle} F(\mathcal{A}(x) - H) dH.$$

But this is the Euclidean Fourier transform of the function $\gamma : H \mapsto e^{\langle \rho,\ H\rangle} F(\mathcal{A}(x) - H)$. When we integrate $e_{\lambda,b_0}(x)\widehat{\psi}(\lambda, b_0)$ over $\mathfrak{a}^\star$ as we must in order to get (5.5), we can use the ordinary Fourier inversion formula (applicable here because γ has compact support and is smooth) and we obtain

$$\int_{\mathfrak{a}^\star} e_{\lambda,b_0}(x)\widehat{\psi}(\lambda, b_0)d\lambda = \gamma(0) = F(\mathcal{A}(x)) = \int_N \psi(ne^{\mathcal{A}(x)} \cdot o)dn,$$

but this is none other than $\int_{\xi(b_0,x)} \psi$. Thus $\delta_{\xi(b,x)}$ and $\mathcal{F}^{-1}\left([\lambda \mapsto e_{\lambda,b_0}(x)] \otimes \delta_{b=b_0}\right)$, both tempered distributions on X (see (A1) below), coincide over $\mathcal{D}(X)$; of course then they do coincide as tempered distributions on X. Taking Fourier transforms proves the lemma. $\square$

Let us come back to the moiré phenomenon (5.1). Remembering our initial wish to swap the integrals in (5.3), let us use the fact that ϑ_ε is rapidly decreasing for every $\varepsilon > 0$ and write

$$\int_X \vartheta_\varepsilon \varphi_\lambda = \int_B \left(\int_X \vartheta_\varepsilon e_{\lambda,b} db\right) = \int_B \widehat{\vartheta_\varepsilon}(\lambda, b)db.$$

The right-hand side is W-invariant in λ (this is obvious from the fact that the left-hand side is, but see also [4], Theorem 2 and [10], Lemma 1.2 p. 200). We can then rewrite the equality as

$$\int_X \vartheta_\varepsilon \varphi_\lambda = \frac{1}{|W|} \sum_{w\in W} \int_B \widehat{\vartheta_\varepsilon}(w\lambda, b)db.$$

Let me now have ε go to zero, so that in the space of tempered distributions on X, ϑ_ε goes to the Dirac distribution over $\xi(b_0, x)$. Because of the continuity properties recalled below in (A1) and (A2), as ε goes to zero, there is a limit in $\mathcal{D}'(\mathfrak{a}^\star/W)$ to the family of distributions given by integration against $\lambda \mapsto |\mathbf{c}(\lambda)|^{-2} \frac{1}{|W|} \int_B \left(\sum\limits_{w\in W} \widehat{\vartheta_\varepsilon}(w\lambda, b)\right) db$ with respect to the measure on $\mathfrak{a}^\star$ inherited from Lebesgue measure. We then see that the distribution associated with integrating against $\lambda \mapsto |\mathbf{c}(\lambda)|^{-2} \int_X \vartheta_\varepsilon \varphi_\lambda$ goes, in $\mathcal{D}'(\mathfrak{a}^\star/W)$, to the distribution $\int_B \mathcal{F}\left(\delta_{\xi(b_0,x)}\right)$ (precisely defined in section 6.2 below), of which the above Lemma says that it is associated with integrating against the almost-everywhere-defined function $\lambda W \mapsto \frac{1}{|W|} \sum\limits_{w\in W} e_{w\lambda,b_0}(x)$ with respect to the Lebesgue-inherited measure.

After a very slight change notation for ϑ in order to revert back from (5.2) to (5.1), we conclude that $\lambda W \mapsto \int_{\xi(b_0,0)} \vartheta(y)\varphi_\lambda^{[y]}(x)\ dy$ goes, in $\mathcal{D}'(\mathfrak{a}^\star/W)$, to the distribution given by integration against $\lambda W \mapsto |\mathbf{c}(\lambda)|^2 \frac{1}{|W|} \sum\limits_{w\in W} e_{w\lambda,b_0}(x)$ with respect to the usual Lebesgue-inherited measure on $\mathfrak{a}^\star/W$: that was the weaker form of (5.1) aimed at in this short chapter, and the proposition is now proven.

6 Tempered distributions on a noncompact symmetric space

6.1 Schwartz functions on X.

Let me write $\mathbf{D}(G)$ for the algebra of left-invariant differential operators on G, and $\bar{\mathbf{D}}(G)$ for the algebra of right-invariant differential operators.

Recall that every element of G can be written as a product k_1ak_2 with k_1, k_2 in K and a in A, and that two such decompositions have their a-part related by the action of an element in the Weyl group. If we set $|g| = |\log(a)|$ (the right-hand-side refers to a Euclidean norm on $\mathfrak{a}$), we can make the following definition: a smooth function f on G is rapidly decreasing (or *Schwartz*) if for every $\ell \in \mathbb{N}$, $L \in \mathbf{D}(G)$ and $R \in \bar{\mathbf{D}}(G)$,

$$\sup_{g\in G} \left|(1+|g|)^{\ell}\Xi(g)^{-1}(LRf)(g)\right| \tag{6.1}$$

is a finite number.

In (6.1), the map Ξ is the spherical function φ_0; in [6], Theorem 3, we find the following estimate :

$$\Xi(g) \leq c(1+|g|)^d e^{-\langle \rho | \log a\rangle}$$

where c is a positive real number and d a nonnegative integer.

The rapidly decreasing functions on G gather in the Schwartz space $\mathcal{S}(G)$; those which are right-invariant under K gather in the Schwartz space $\mathcal{S}(X)$. The quantities (6.1) provide natural seminorms turning $\mathcal{S}(X)$ into a Fréchet space; I write $\mathcal{S}'(X)$ for the topological dual $\mathcal{S}(X)$, the space of *tempered distributions* on X.

6.2 Schwartz functions on $\mathfrak{a}^\star \times K/M$ and continuity of the Fourier transform.

When taking the Helgason-Fourier transform of a function in $\mathcal{S}(X)$, we get a smooth function on $\mathfrak{a}^\star \times K/M$ which satisfies ([10], chap. 3, thm 1.10):

$$\text{For each } P \in \mathbb{R}[X,Y] \text{ and every } \ell \in \mathbb{N}, \quad \sup_{\lambda,b}\left|(1+|\lambda|)^{\ell}\left(P(\Delta_{K/M},\Delta_{\mathfrak{a}^\star})\cdot g\right)(\lambda,b)\right| < \infty \tag{6.2}$$

with $\Delta_{K/M}$, $\Delta_{\mathfrak{a}^\star}$ the Laplace-Beltrami operators on B and $\mathfrak{a}^\star$.

Let us write $\mathcal{S}(\mathfrak{a}^\star \times K/M)$ for the space of smooth functions on $\mathfrak{a}^\star \times B$ for which (6.2) holds; as before it comes with natural seminorms which make it a Fréchet space; taking Fourier-Helgason transforms of course defines a continuous, injective map from $\mathcal{S}(X)$ into $\mathcal{S}(\mathfrak{a}^\star \times K/M)$.

Harish-Chandra and Helgason proved that this map defines a homeomorphism between the subspaces gathering the $K-$invariants in both spaces ([10], th. 1.17; see also Anker [2]). Eguchi [3, 4] proved (together with Okamoto) that the Fourier transform of an element f of $\mathcal{S}(\mathfrak{a}^\star \times K/M)$ satisfies some form of Weyl-group invariance (see [4], Theorem 2): the averages over B of $b \mapsto \widehat{f}(\lambda,b)$ $b \mapsto \widehat{f}(w\lambda,b)$ coincide for every w in W. Writing $\mathcal{S}(\mathfrak{a}^\star \times K/M)_W$ for the space of Schwartz functions on $\mathfrak{a}^\star \times K/M$ satisfying that Weyl-group invariance condition, Eguchi and Okamoto proved that $\mathcal{F}$ induces a homeomorphism between $\mathcal{S}(X)$ and $\mathcal{S}(\mathfrak{a}^\star \times K/M)_W$.

Now, define the space $\mathcal{S}'(\mathfrak{a}^\star \times K/M)$ of tempered distributions on $\mathfrak{a}^\star \times K/M$ as the topological dual of $\mathcal{S}(\mathfrak{a} \times K/M)_W$, the space $\mathcal{S}'(X)$ of tempered distributions on X as the topological dual of $\mathcal{S}(X)$, and if T is a tempered distribution on X, define $\hat{T}$ as the distribution $\psi = \hat{\varphi} \in \mathcal{S}(\mathfrak{a}^\star \times B)_W \mapsto \langle T \mid \varphi\rangle$ on $\mathfrak{a}^\star \times B$. Then of course:

$$T \mapsto \hat{T} \text{ defines a homeomorphism between } \mathcal{S}'(X) \text{ and } \mathcal{S}'(\mathfrak{a}^\star \times K/M). \tag{A1}$$

To complete our list of disribution-theory-based ingredients for section 5, let me record the following remark: if T is an element of $\mathcal{S}'(\mathfrak{a} \times K/M)$, one can define a distribution U

on $\mathfrak{a}^\star$ ("the integral of T over K/M") by setting, for ζ in $\mathcal{D}(\mathfrak{a}^\star/W)$, $\langle U \mid \zeta\rangle = \langle T \mid \tilde{\zeta} \otimes 1_B\rangle$ (where $\tilde{\zeta}$ is the inflation of ζ to $\mathfrak{a}^\star$). Writing $\int_B T$ for the distribution U, we then of course have:

$$\text{The map } T \mapsto \int_B T \text{ is continuous as a map from } \mathcal{S}'(\mathfrak{a}^\star \times K/M) \text{ to } \mathcal{D}'(\mathfrak{a}^\star/W). \tag{A2}$$

I should point out here that an almost-everywhere defined and bounded function u on $\mathfrak{a}^\star$ defines an element of $\mathcal{S}'(\mathfrak{a}^\star \times K/M)$, but that given the form of the Plancherel formula for the Helgason-Fourier transform (see [10], p. 203), if the above definition of the Fourier transform of tempered distributions is to extend the Helgason-Fourier transform of smooth and rapidly decreasing functions, the distribution should be given by integration against $|\mathbf{c}(\lambda)|^{-2}u(\lambda, b)d\lambda \otimes db$ rather than against $u(\lambda, b)d\lambda \otimes db$.

Bibliography

[1] I. Amidror, *The theory of the moiré phenomenon*, Computational Imaging and Vision (2000), vol. 15.

[2] J-Ph. Anker , *The spherical Fourier transform of rapidly decreasing functions*, J. Funct. Anal. **96-2** (1991), pp. 331–349.

[3] M. Eguchi, *Asymptotics of Eisenstein Integrals* Journal of Functional Analysis **34** (1979), pp. 167-216

[4] M. Eguchi & K. Okamoto, *The Fourier Transform of the Schwartz Space on a Symmetric Space*, Proc. Japan Acad. Ser. A **53 -7** (1977), pp. 237-241

[5] Harish-Chandra, *Spherical functions on a semisimple Lie group*, Proc Natl Acad Sci U S A. **43(5)** (1957), pp. 408–409.

[6] Harish-Chandra, *Spherical functions on a semisimple Lie group. I*, American Journal of Mathematics **80** (1958), pp. 241–310

[7] J. Harthong, *Le moiré.* Publications de l'IRMA, Strasbourg, vol. 93 (1981).

[8] S. Helgason, *Differential geometry, Lie groups, and symmetric spaces.* Academic press, vol. 80 (1979).

[9] S. Helgason, *Groups & Geometric Analysis: Radon Transforms, Invariant Differential Operators and Spherical Functions.* Academic press (1984).

[10] S. Helgason, *Geometric analysis on symmetric spaces*, 2nd Ed. American Mathematical Society, vol. 39 (2008).

[11] D. Hubel, *Eye, Brain and Vision.* Scientific American Library (1995).

[12] D. H. Hubel and T. N. Wiesel, *Receptive fields, binocular interaction and functional architecture in the cat's visual cortex*, Journal of Physiology **160-1** (1962), pp. 106-154.

[13] D. Hubel, *Transformation of information in the cat's visual system*, Proceedings of the International Union of Physiological Sciences, XXII International Congress, Leiden (1962).

Chapter 2

Monochromaticity of orientation maps in V1 implies minimum variance for hypercolumn size

Contents

Abstract. In the primary visual cortex of many mammals, the processing of sensory information involves recognizing stimuli orientations. The repartition of preferred orientations of neurons in some areas is remarkable : a repetitive, non- periodic, layout. This repetitive pattern is understood to be fundamental for basic non-local aspects of vision, like the perception of contours, but important questions remain about its development and function.

We focus here on Gaussian Random Fields, which provide a good description of the initial stage of orientation map development and, in spite of shortcomings we will recall, a computable framework for discussing general principles underlying the geometry of mature maps. We discuss the relationship between the notion of column spacing and the structure of correlation spectra ; we prove formulae for the mean value and variance of column spacing, and use numerical analysis of exact analytic formulae to study the variance. Referring to studies by Wolf, Geisel, Kaschube, Schnabel and coworkers, we also show that spectral thinness is not an essential ingredient to obtain a pinwheel density of π, whereas it appears as a signature of Euclidean symmetry. The minimum variance property associated to thin spectra could be useful for information processing, provide optimal modularity for V1 hypercolumns, and be a first step towards a mathematical defintion of hypercolumns. A measurement of this property in real maps is in principle possible, and comparison with the results in our paper could help establish a role of our minimum variance hypothesis in the development process.

1 Introduction

Neurons in the primary visual cortex (V1, V2) of mammals have stronger responses to stimuli that have a specific orientation [1, 2, 3]. In many species including primates and carnivores (but no rodent, even though some of them have rather elaborated vision [4, 11]), these orientation preferences are arranged in an ordered map along the cortical surface. Moving orthogonally to the cortical surface, one meets neurons with the same orientation preference; travelling along the cortical surface, however, reveals a striking arrangment in smooth, quasi- periodic maps, with singular points known as *pinwheels* where all orientations are present [5, 6, 7], see fig. 1. All theses orientation maps look similar, even in distantly related species [12, 11]; the main difference between any two orientation preference maps (OPM) seems to be a matter of global scaling.

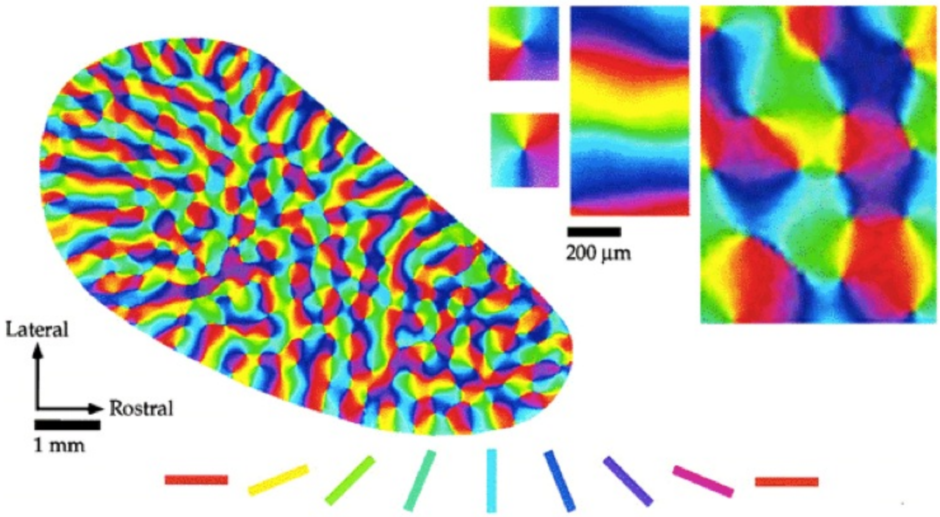


Figure 1: **Layout of orientation preferences in the visual cortex of a tree shrew** (modified from Bosking et al [35]). Here orientation preference is color- coded (for instance neurons in blue regions are more sensitive to vertical stimuli). Maps of sensivity to different stimulus angles were obtained by optical imaging; summing these with appropriate complex phases yields figure 1: see Swindale [14]. In particular, at singular points (pinwheels), all orientations meet (see the upper right corner); for a fine- scale experimental study of the neighborhood of such points, see [7].

The common design has very precise and beautiful geometrical properties, and universal quantitative properties of these cortical maps have recently been uncovered: for instance, a density of singular points close to π has been observed [12], see below. However the exact functional advantage of this geometrical arrangment in hypercolumns remains unclear [11, 14, 15, 16, 17]. What is more, the functional principles underlying the observed properties of orientation maps are still in debate; in particular, it is often thought that a pinwheel density of π has to do with monochromaticity (existence of a critical wavelength in the correlation spectrum) of the cortical map. The aim of this short paper is to clarify the role of the monochromaticity, or spectral thinness, condition, using the simplified mathematical framework of Gaussian Random Fields with symmetry properties.

Our first few remarks (section 2.1) are included for clarification purposes: we first give an intrinstic definition of the column spacing in these fields, then discuss the intervention of spectral thinness in theoretical and experimental results related to pinwheel densities. Then (section 2.2) we introduce the “minimum variance" property in our title, to help discuss the quasi- periodicity in the map and to try to understand better the notion of cortical hypercolumn. In the concluding Discussion, we also try to clarify the relevance of this property for the development of real maps and formulate a simple test for our hypothesis that it is indeed relevant.

Many models for the development of orientation maps have been put forward [18, 19, 11]; they address such important issues as the role of self- organization, or of interactions between orientation and other parameters of the receptive profiles [16]. In this short note, we focus on a mathematical computable framework in which geometrical properties can be discussed with full proofs, and whose quantitative properties can now be compared with those of experimental maps. While we thus put the focus on the geometry of theoretical maps rather than on the most realistic developmental scenarios, we try to relate this geometry to organizing principles, *viz.* information maximization and perceptual invariance, which are relevant for discussing real maps. In a mathematitical setting, these principles can be enforced through explicit randomness and invariance structures.
Wolf, Geisel, Kaschube and coworkers [20, 21, 22, 12] have described a wide class of probabilistic models for the development of orientation preference maps. In all these models (and in our discussion) the cortical surface is identified with the plane $\mathbb{R}^2$, and the orientation preference of neurons at a point x is given by (half) the argument of a complex number $\mathbf{z}(x)$; one adds the important requirement that the map $x \mapsto \mathbf{z}(x)$ be continuous (this is realistic enough if the modulus $|z(x)|$ stands for something like the orientation selectivity of the neurons at x, see [14, 6, 27]). Pinwheel centers thus correspond to zeroes of $\mathbf{z}$.

A starting point for describing orientation maps in these models, one which we will retain in this note, is the following general principle: we should treat $\mathbf{z}$ as a random field, so at each point x, the complex number $\mathbf{z}(x)$ as a random variable.

Even without considering development, it is reasonable to introduce randomness, to take into account inter-individual variability. But of the statistical properties of zero-set of general random fields, our understanding is that present- day mathematics can say very little [25]; only for very specific subclasses of random fields are precise mathematical theorems available. The most important of those is the class of *Gaussian Random fields* [24, 25, 26] − a random field $\mathbf{z}$ is Gaussian when all joint laws for $(\mathbf{z}(x_1), ...\mathbf{z}(x_n)) \in \mathbb{C}^n$ are gaussian random variables.

If the map $\mathbf{z}$ arises from an unknown initial state and if the development features a stochastic differential equation, taking into account activity-dependent fluctuations and noise, the Gaussian hypothesis is very natural for the early stages of visual map development (see [20, 22, 28]). In the most precise and recent development models by Wolf, Geisel, Kaschube and others [39, 12, 16], it is however only the initial stage that turns out to be well- represented by a Gaussian field: upon introducing long- range interactions in the integral kernel of the stochastic differential equation representing the refinment of cortical circuitry, the Gaussian character of the field must be assumed to break down when the nonlinearities become significant, and the stationary states of the dynamics which represent mature maps cannot be expected to be Gaussian states. We shall comment on this briefly in section 2.1.3, and come back to it in the Discussion.

In spite of this, we shall stick to the geometry of maps sampled from Gaussian Random Fields (GRFs) in this short paper. We have several reasons for doing this. A first remark is that that a better understanding of maps sampled from them can be helpful in understanding the general principles underlying more realistic models, or helpful in suggesting some such principles. A second remark is that with the naked eye, it is difficult to see any difference between some maps sampled from GRFs and actual visual maps (see figure 2), and that there is a striking likeness between some theorems on GRFs and some properties measured in V1. A third is that precise mathematical results on GRFs can be used for testing how close this likeness is, and to make the relationship between GRFs and mature V1 maps clearer.

Wolf and Geisel add a requirement of *Euclidean invariance* on their stochastic differential equation, so that if the samples from a GRF are to be thought of as providing (early or mature) cortical maps, the field should be homogeneous (*i.e.* insensitive, as a random field, to a global translation $x \mapsto x+a$), isotropic (insensitive to a global rotation, $x \mapsto \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix} x$) and centered (insensitive to a global shift of all orientation preferences, changing the value $\mathbf{z}(x)$ at each x to $e^{i\theta}\mathbf{z}(x)$). Here again, looking at mature maps, geometrical invariance is a natural requirement for perceptual function; so we shall assume that the GRF $\mathbf{z}$ is centered, homogeneous and isotropic [25, 29]. Note that of course, this invariance requirement cannot be formulated in a non-probabilistic setting (a deterministic map from $\mathbb{R}^2$ to $\mathbb{C}$ cannot be homogeneous without being constant).

It actually turns out that these two mathematical constraints (gaussian field statistics and symmetry properties) are strong enough to generate realistic- looking maps, with global quasi- periodicity. Quite strikingly, it has been observed [30, 28] that *one needs only add a spectral thinness condition* to obtain maps that seem to have the right qualitative (a hyper columnar, quasi- periodic organization) *and quantitative* properties (a value of π for pinwheel density). These mathematical features stand out among theoretical models for orientation maps as producing a nice quasiperiodicity, with rougly repetitive "hypercolumns" of about the same size that have the same structure, as opposed to a strictly periodic crystal- like arrangment (see [16], compare [41]). The aim of this short note is to clarify the importance of this spectral thinness condition for getting a quasi- periodic “hypercolumnar" arrangment on the one hand, a pinwheel density of π on the other.

Before we give results about homogeneous and isotropic GRFs, let us mention that the quantitative properties of the common design which have been observed by Kaschube et al. [12] also include mean values for three kinds of nearest neighbour distance and for two parameters representing the variability of pinwheel density as a function of subregion size; evaluating these mean values in the mathematical setting of random fields, even in the oversimplified case of invariant GRFs, is a difficult mathematical problem which is beyond the author’s strengths at present. So in this short note, we shall focus on the existence of a precise hypercolumn size and a well- defined pinwheel density in the common design, and refrain from examining the other important statistics.

2 Results

2.1 Two remarks on Gaussian Random Fields with Euclidean symmetry

2.1.1. Let us first formulate the spectral thinness condition more precisely: in an invariant GRF, the correlation $C(x, y)$ between orientations at x and y depends only on $\|x - y\|$. Let us turn to its Fourier transform, or rather to the Fourier components of the map $\Gamma : \mathbb{R}^2 \to \mathbb{C}$ such that $C(x, y) = \Gamma(x - y)$. For an invariant gaussian field, specifying Γ does determine the field; what is more, there is a unique measure P on $\mathbb{R}^+$ such that

$$\Gamma(\tau) = \int_{R>0} \Gamma_R(\tau)\ dP(R) \tag{2.1}$$

where, for fixed $R > 0$, the map Γ_R is[1] $\tau \mapsto \int_{\mathbb{S}^1} e^{iR\vec{u}\cdot\tau} d\vec{u}$.

Now, correlations on real cortical maps can be measured and the spectrum of Γ can be inferred [28]; data obtained by optical imaging reveals that the spectral measure P is concentrated on an annulus ([28], p. 100, see also [30]): this means that there is a *dominant wavelength* Λ_0, such that the measure P concentrates around $R_0 = \frac{2\pi}{\Lambda_0}$.

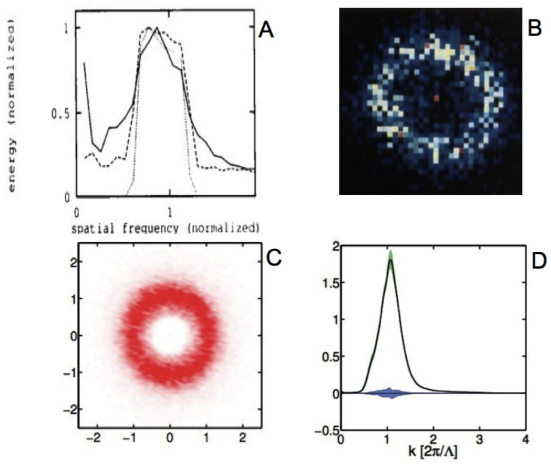


Figure 2: **Correlation spectra of orientation maps in macaque and tree shrew V1**. A and B are from Niebur and Worgotter's 1994 paper [30]: in A, the solid and dashed lines are spectra obtained by two different methods (direct measurement of correlations and Fourier analysis) from an experimental map obtained by Blasel in macaque monkey, the power spectrum of which is displayed on B. Images C and D are from Schnabel's 2008 thesis [28], p. 104. Methods for obtaining C and D from measurements on Tree Shrews are explained precisely by Schnabel in [28], sections 5.3 and 5.4. The green- and blue-shaded regions code for boostrap confidence interval and 5% significance level, respectively. The power spectrum in D has standard deviation around 0.2 in the unit displayed on the horizontal axis and determined by the location of the maximum; the mean and quadratic wavenumbers in this spectrum are in the intervals [1.05, 1.10] and [1.18,1.23], respectively.

Correlation spectra of real V1 maps, first discussed in [30], have been measured precisely by Schnabel in tree shrews [28]: (see [28], p. 104, fig. 5.6(d) is reproduced on figure 2). The spectral measure P has a nicely peaked shape, and the very clear location of the peak

1. The measure on $\mathbb{S}^1$ used in this formula has total mass one.

is used as the dominant wavelength Λ.; see figure 2. From Schnabel's data we evaluate the standard deviation in P to be about 0.2Λ (caution: here P is a real correlation spectrum, not the spectral density of a GRF).

Although this is far from being an infinitely thin spectrum, it is not absurd to look at the extreme situation where we impose the spectral thinness to be zero. Figure 3 shows a map sampled from a *monochromatic* invariant GRF, in which Γ is one of the maps Γ_R of the previous paragraph, in other words the inverse Fourier transform of the Dirac distribution $\delta(R - R_0)$ on a circle: monochromatic, or almost monochromatic, invariant GRFs yield quite realistic-looking maps, at least to the naked eye.

This thinness hypothesis certainly has to do with the existence of a precise scale in the map, that is, with the "hyper columnar" organization. In all existing theoretical studies that we know of, spectral thinness is introduced *a priori* into the equations precisely in order to obtain a repetitive pattern in the model orientation maps. For instance, in the very successful long- range interaction model of Wolf and al. [39, 12], the linear part of the stochastic differential equation for map development features a Swift- Hohenberg operator in which a characteristic wavelength is imposed. The "typical spacing" between iso-orientation domains is then defined as that which corresponds to the mean wavenumber in the power spectrum:

$$\frac{2\pi}{\Lambda_{mean}} := \int k dP(k). \tag{2.2}$$

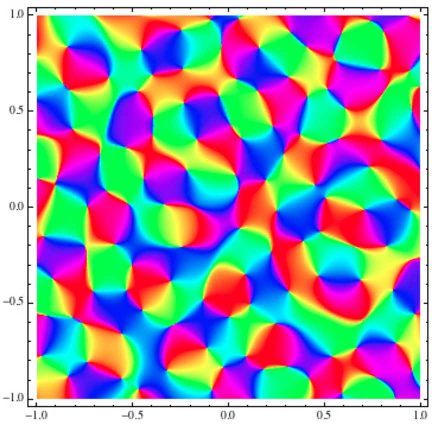


Figure 3: **Computer- generated map, sampled from a monochromatic field**. This figure shows an orientation map which we have drawn from a simulated Invariant Gaussian Random Field with circular power spectrum. We used 100 plane waves with frequency vectors at the vertices of a regular polygon inscribed in a circle, and random Gaussian weights (see the additional documentation); with respect to the unit of length displayed on the $x-$ and $y-$ axes, the wavelength of the generating plane waves is $1/3$.

2.1.2. It is reasonable, both intuitively and practically, to expect that Λ_{mean} gives the mean local period between iso- orientation domains. For reasonable bell- shaped power spectra, Λ_{mean} is in addition quite close to the location of the peak in the spectrum, which very obviously corresponds to the “dominant frequency" in the powerspectrum and is quite straightforward to measure. But from a mathematical point of view, there is a paradox here.

For Gaussian fields, it is natural to try to clarify this and write down an intrinstic definition of the mean column spacing in terms of the probabilistic structure of the field. The paradox is that then the natural scale to use turns out to be different from Λ_{mean}, and the difference is appreciable in measured spectra. We are going to show presently that in an invariant gaussian random field, the typical spacing turns out to be the wavelength Λ_{sq} corresponding to the quadratic mean wavenumber:

$$\frac{2\pi}{\Lambda_{sq}} := \sqrt{\int k^2 dP(k)}, \tag{2.3}$$

which coincides with Λ_{mean} if and only if the field is monochromatic.

Using Schnabel's data to evaluate the corresponding wavelengths in real maps, we find that the quotient between Λ_{sq} and Λ_{mean} is about 1.1, and they are within 15% of each other. So using one rather than the other does have an importance.

To justify our claim that Λ_{sq} is a good intrinstic way to define the column spacing in an invariant Gaussian field, let us consider a fixed value of orientation, say the vertical. Let us draw any line D on the plane and look for places on D where this orientation is represented, which means that the real part of $\mathbf{z}$ vanishes. Now if $\mathbf{z}$ is an Euclidean-invariant standard gaussian field, $\mathfrak{Re}(\mathbf{z})|_D$ is a translation- invariant gaussian field on the real line D. From the celebrated formula of Kac and Rice we can then deduce the typical spacing between its zeroes, and this yields the following theorem :

Result 1: pick any line segment J of length ℓ on the plane and any orientation $\theta_0 \in \mathbb{S}^1$. Write $\mathcal{N}_{J,\theta_0}$ for the random variable recording the number of points on J where the gaussian field $\mathbf{z}$ provides an orientation θ_0. Then

$$\mathbb{E}[\mathcal{N}_{J,\theta_0}] = \frac{\ell}{\Lambda_{sq.}}$$

Indeed, let us write Φ for $\mathfrak{Re}(\mathbf{z})|_D$, viewed as a stationary Gaussian field on the real line, G for its covariance function, and $\mathcal{G}$ for the covariance function of $\mathfrak{Re}(\mathbf{z})$ viewed as a homogeneous and isotropic random field on $\mathbb{R}^2$. The arguments leading up to the statement of Result 1 and the Kac- Rice formula which is recalled in the additional documentation prove that $\mathbb{E}[\mathcal{N}_{J,\theta_0}] = \ell \cdot \frac{\sqrt{\lambda}}{\pi}$, where $\lambda = \mathbb{E}\left[\Phi'(0)^2\right]$. But $\mathbb{E}\left[\Phi'(0)^2\right] = \partial_{x_1}\partial_{x_2}\mathbb{E}\left[\Phi(x_1)\Phi(x_2)\right]\big|_{x_1=x_2=0}$, and this is $\partial_x\partial_y G(x-y)\big|_{x=y=0} = -G''(0)$. To complete the proof we need to calculate this.

Now, in view of the Euclidean invariance of $\mathfrak{Re}(\mathbf{z})$, we know that $G''(0)$ is half the value of $\Delta\mathcal{G}$ at zero. To evaluate this quantity, we use the spectral decomposition of $\mathcal{G}$: it reads $\mathcal{G} = \int_{R>0} \mathcal{G}_R dP(R)$, where $\mathcal{G}_R$ is the covariance function of a real- valued monochromatic invariant field on $\mathbb{R}^2$, hence is equal to $\frac{1}{2}\Gamma_R$ (recall that Γ_R was defined in equation (1), and is real- valued). Now, Γ_R satisfies the Helmholtz equation $\Delta(\Gamma_R) = -R^2\Gamma_R$, and in addition $\Gamma_R(0)$ is equal to one, so $\mathcal{G}_R(0)$ is equal to one- half. We conclude that $G''(0)$ is equal to $-\frac{1}{4}\int_{R>0} R^2 dP(R) = \frac{\pi^2}{\Lambda_{sq}^2}$. This completes the proof of Result 1.

Let us now comment on this result. It means that repetitions of θ_0 occur in the mean every $\Lambda_{sq.}$. Of course this is very close to Λ_{mean} when the support of the power spectrum is contained in a thin enough annulus (if the width of such an annulus is less than a fifth of its radius, Λ_{mean} and Λ_{sq} are within 3 % of each other). But in general, it is obvious from

Jensen's inequality that $\Lambda_{mean} \geq \Lambda_{sq.}$, with equality if and only if the field is monochromatic. In real maps, there is an appreciable difference between Λ_{mean} and Λ_{sq} as we saw.

2.1.3. Let us turn now to pinwheel densities; we would like to comment on a beautiful theoretical finding by Wolf and Geisel and related experimental findings by Kaschube, Schnabel and others. We feel we should be very clear here and insist that this subsection is a comment on work by Wolf, Geisel, Kaschube, Schnabel and others; if we include the upcoming discussion it is to clarify the role of the spectral thinness condition in the proof of their result, and we seize the opportunity to comment on this work's theoretical significance.

If a wavelength Λ is fixed, the pinwheel density d_Λ in a (real or theoretical) map is the mean number of singularities in an area Λ^2. In the experimental studies of Kaschube et al. [12] and Schnabel [28], the wavelength used is obtained with two algorithms, one which localizes the maximum in the power spectrum, and one which averages local periods obtained by wavelet analysis. These two algorithms give approximately the same result, say Λ_{exp}, and pinwheel densities are scaled relatively to this Λ_{exp}: a very striking experimental result is obtained by Kaschube's group, namely

$$d_{\Lambda_{exp}} = \text{ mean number of pinwheels in a region of area } \Lambda_{exp}^2 \simeq \pi \pm 2\%. \tag{2.4}$$

On the other hand, in an *invariant* gaussian random field, expectations for pinwheel densities may be calculated using generalizations of the formula of Kac and Rice. This calculation has been conducted by Wolf and Geisel [20, 22], Berry and Dennis [31]; recent progress on the mathematical formulation of the Kac-Rice formula makes it possible to write down new proofs [26, 32], as we shall see presently. The value of π occurs very encouragingly here, too:

Theorem (Wolf and Geisel [22], Berry and Dennis [31], see also [26, 32]) : let us write $\mathcal{P}_{\mathcal{A}}$ for the random variable recording the number of zeroes of the gaussian field $\mathbf{z}$ in a region $\mathcal{A}$, and $|\mathcal{A}|$ for the euclidean area of $\mathcal{A}$. Then

$$\mathbb{E}[\mathcal{P}_{\mathcal{A}}] = \frac{\pi}{\Lambda_{sq.}^2}|\mathcal{A}|.$$

We think it can be of interest for readers of this journal that we include a proof of this result here. We would like to say very clearly that the discovery of this result is due to Wolf and Geisel on the one hand, and independently to Berry and Dennis in the monochromatic case. In [26], Azais and Wschebor gave a mathematically complete statement of a Kac-Rice- type formula, and recently Azais, Wschebor and Leon used it (following Berry and Dennis) to give a mathematically complete proof of the above theorem, though they wrote down the details only in case $\mathbf{z}$ is monochromatic [32]. It is for the reader's convenience, and because the focus of this short note is with non- monochromatic fields, that we recall their arguments here.

Azais and Wschebor's theorem (Theorem 6.2 in [26]), in the particular case of a smooth reduced gaussian field, is the following equality :

$$\mathbb{E}(\mathcal{P}_{\mathcal{A}}) = \frac{1}{2\pi}\int_{\mathcal{A}} \mathbb{E}\left\{|\det\, d\mathbf{z}(p)| \mid \mathbf{z}(p) = 0\right\} dp.$$

Here the integral is with respect to Lebesgue measure on $\mathbb{R}^2$, and the integrand is a conditional expectation.

To evaluate this, one should first note that $\mathbf{z}$ has constant variance, and an immediate consequence is that for each p, the random variables $\mathbf{z}(p)$ is independent from the random variable recording the value of the derivative of the real part (resp. the imaginary part) of $\mathbf{z}$ at p. So the random variables $|\det d\mathbf{z}(p)|$ and $\mathbf{z}(p)$ are actually independent at each p, and we can remove the conditioning in the formula. Now at each p, $d\mathbf{z}(p)$ is a 2×2 matrix whose columns, $C_1(p) := \begin{pmatrix} (\partial_x \mathfrak{Re}(\mathbf{z}))(p) \\ (\partial_y \mathfrak{Re}(\mathbf{z}))(p) \end{pmatrix}$ and $C_2(p) := \begin{pmatrix} (\partial_x \mathfrak{Im}(\mathbf{z}))(p) \\ (\partial_y \mathfrak{Im}(\mathbf{z}))(p) \end{pmatrix}$, are independent gaussian vectors (see [25], section 1.4 and chapter 5). Because $\mathbf{z}$ has Euclidean symmetry, $C_1(p)$ and $C_2(p)$ have zero mean and the same variance, say V_p, as $(\partial_x \mathfrak{Re}(\mathbf{z}))(p)$. But $|\det d\mathbf{z}(p)|$ is the area of the parallelogram generated by $C_1(p)$ and $C_2(p)$, and the "base times height" formula says this area is the product of $\|C_1(p)\|$ with the norm of the projection of $C_2(p)$ on the line orthogonal to $C_1(p)$. The expectation of $\|C_1(p)\|$, a "chi- square" random variable, is $2\sqrt{V_p}$ and that of the norm of the projection of $C_2(p)$ on any fixed line is $\sqrt{V_p}$; since both columns are independent, we can conclude that

$$\mathbb{E}(\mathcal{A}) = \frac{1}{\pi} \int_{\mathcal{A}} V_p dp = \frac{|\mathcal{A}|}{\pi} V_0$$

(the last equality is because $\mathbf{z}$ and all its derivatives are stationary fields). Now we need to evaluate $V_0 = \mathbb{E}\left\{(\partial_x \mathfrak{Re}\mathbf{z})(0)^2\right\}$. But this quantity already appeared in the proof of Result 1, it was labelled λ there. So we already proved that it is equal to $\frac{\pi^2}{\Lambda_{sq}^2}$, and this concludes the proof of Wolf and Geisel's Theorem.

From this theorem Wolf, Geisel and others deduce that $d_{\Lambda_{mean}} \geq \pi$, and it is in this form that the Theorem is discussed. However, we have seen that $d_{\Lambda_{sq}}$, which is equal to π whatever the spectrum, is a rather more natural theoretical counterpart to $d_{\Lambda_{exp}}$. If we drop the focus away from Λ_{mean} to bring Λ_{sq} to the front, we obtain from Result 1 the following reformulation of Wolf and Geisel's theorem :

Result 2 : Write Δ for the typical distance between iso- orientation domains, as expressed by Result 1, and η for the value $\frac{\mathbb{E}[\mathcal{P}_{\mathcal{A}}]}{|\mathcal{A}|}$ of pinwheel density. Then

$$\eta = \frac{\pi}{\Delta^2}. \tag{2.5}$$

There are two simple consequences of Wolf and Geisel's finding which we would like to bring to our reader's attention.

The first is that the pinwheel density of π observed in experiments is scaled with respect to Λ_{exp}, and not with respect to Λ_{sq}. Using Schnabel's data, we can evaluate the $d_{\Lambda_{sq}}$ of real maps, and as Λ_{sq} is about $0.82\Lambda_{\text{exp}}$ in Schnabel's data, $d_{\Lambda_{sq}}$ strongly departs from π in real maps. Since it would be exactly π in maps sampled from GRFs, one consequence of the work in [22, 28, 12] is the following

Corollary : The pinwheel density of observed mature maps is actually *incompatible* with that of maps sampled from invariant Gaussian Fields.

This fact is quite apparent in the work by Wolf, Geisel, Kaschube and coworkers, but since we focused on GRFs in this short note we felt it was useful to recall this as clearly as possible.

Our second remark is that in the reformulation stated as Result 2 here, there is no longer any spectral thinness condition. In other words, *when we consider maps sampled from Gaussian Random Fields, a pinwheel density of π is a numerical signature of the* *fact that the field has Euclidean symmetry.* Result 2 thus shows that when one considers invariant GRFs, average pinwheel density and monochromaticity are independent features.

Because invariant GRFs have ergodicity properties, an ensemble average such as that in Result 2 can be evaluated on an individual sample map; one can thus consider a single output of the GRF $\mathbf{z}$ and proceed to quantitative measurements on it to determine whether the probability distribution of $\mathbf{z}$ has Euclidean symmetry. Very remarkable since no single output can have Euclidean symmetry !

To conclude this subsection, let us recall that Results 1 and 2 say nothing of map ensembles that do not have Gaussian statistics, and in particular nothing of the geometry of real maps; they certainly do not mean that the defintion of Λ_{exp} used in experiments is faulty, but were simply aimed at disentangling monochromaticity from other geometrical principles in the simplified setting of GRFs. To illustrate the fact that our results are not incompatible with the definition of Λ_{exp} used in experiments, let us note that of the two methods used by Kaschube and coworkers to determine Λ_{exp}, one (the averaging of local wavelet- evaluated spacings) provides a definition of column spacing similar to that which we used in Result 1, and the other (looking for the peak in the powerspectrum) gives an appreciably different result from Λ_{sq} as we recalled. The fact that Kaschube et al. observe the two algorithms to give very close results in real maps does not go against Result 1, but rather can be seen as another argument, this time Result- 1- based, against GRFs representing mature maps. The measurement of pinwheel density, equation (4), furthermore indicates that development seems to keep Result 2 true at the mature stage. We shall come back to this in the Discussion.

2.2 The variance of column spacings

Results 1- 2 show that for Gaussian Random Fields, the existence of a pinwheel density of π is independent of the monochromaticity condition. We evaluated the expected value of the column spacing in an invariant GRF in Result 1, and we now turn to its variance. There are several reasons why it should be interesting to establish rigorously that spectral thinness provides a low variance.

A first one is the search for a mathematically well- defined counterpart to the statement, visually obvious, that orientation maps are “quasiperiodic". Most mathematical definitions of quasiperiodicity (like those which follow Harald Bohr [42]) are not very well- suited to discussing V1 maps, and we feel that the meaning of the word is, in the case of V1 maps, well- conveyed by the property we will demonstrate. While it is intuitively obvious that a “nice quasiperiodicity" should come with spectral thinness, as we shall see it is mathematically non trivial.

A second reason to look at the variance is to try to understand better the concept of “cortical hypercolumn", due to Hubel and Wiesel, which is crucial to discussions of the functional architecture of V1. Neurons in V1 are sensitive to a variety of local features of the visual scene, and a hypercolumn gathers neurons whose receptive profiles span the possible local features (note that there is no well- defined division of V1 in hypercolumns, but an infinity of possible partitionings). In studies related to the local geometry of V1 maps, once a definition for the column spacing Λ has been chosen, one is led (as in [21, 23, 12]) to define the area of a hypercolumn as Λ^2. Here we put the focus on the orientation map only; but even then is thus legitimate to wonder whether in a domain of area Λ^2, each orientation is represented at least once. Note that a value of π for pinwheel density can guarantee this if one establishes that the density also has a small variance; here however we are not going to evaluate this variance, which is possible in principle [32] but not easy, and simply focus on column spacing. This is a first step in trying to check that the internal structure of domains with area Λ^2 is somewhat constant, as suggested by the results on pinwheel density

Let us add that from the point of view of information processing, it is not unnatural to expect a low variance for hypercolumn size. It is known that the behaviour of many neurons in the central nervous system depends on the statistical properties in the distributions of spikes reaching them, and not only on the average activity. These statistical characteristics depend on physiology of course, but also on the information being vehicled. Now, vision is an active process: the eye moves ceaselessly and a given object or contour is processed by many regions of V1 in a relatively short time. For a neuron receiving inputs from V1, a low variance for hypercolumn size should help make the distribution of received informations more uniform (with minimum bias for a given orientation). This would be in harmony with a general principle at work in the central nervous system, that of maximizing mutual information, which on the sensory side corresponds to a maximum of discimination (and Fisher information, see [44]) and on the motor side to what has been called the "minimum variance principle", for instance in the study of ocular saccades or arm movements [43].

So we will now consider the variance $\mathbb{V}[\mathcal{N}_{J,\theta_0}]$ of the previous random variable. We will show that it reaches a minimum when the spectrum is a pure circle. Now, evaluating this variance is surprisingly difficult, even though there is an explicit formula, namely

Theorem (Cramer and Leadbetter, see [33]) : In the setting of Result 1, write $G : \mathbb{R} \to \mathbb{R}$ for the covariance function of $\mathfrak{Re}(\mathbf{z})|_D$ and $M_{33}(\tau)$, $M_{44}(\tau)$ the cofactors of the $(3,3)^{rd}$ and $(3,4)^{th}$ entries in the matrix

$$\begin{pmatrix} 1 & G(\tau) & 0 & G'(\tau) \\ G(\tau) & 1 & -G'(\tau) & 0 \\ 0 & -G'(\tau) & -G''(0) & -G''(\tau) \\ G'(\tau) & 0 & -G''(\tau) & -G''(0) \end{pmatrix}.$$

Then

$$\mathbb{V}[\mathcal{N}_{J,\theta_0}] = \frac{\pi\ell}{\Lambda_{sq}} - \left(\frac{\pi\ell}{\Lambda_{sq}}\right)^2 + \frac{2}{\pi^2}\int_0^\ell (\ell-\tau)\frac{\sqrt{M_{33}(\tau)^2 - M_{34}(\tau)^2}}{(1-G(\tau)^2)^{3/2}} \tag{2.6}$$

$$\left[1 + \frac{M_{34}(\tau)}{\sqrt{M_{33}(\tau)^2 - M_{34}(\tau)^2}} \arctan\left(\frac{M_{34}(\tau)}{\sqrt{M_{33}(\tau)^2 - M_{34}(\tau)^2}}\right)\right] d\tau.$$

Recall here that

$$G(\tau) = \frac{1}{4\pi}\int_{R>0}\left(\int_0^{2\pi}\cos(R\tau\cos(\vartheta))d\vartheta\right)P(R)dR : \tag{2.7}$$

this $G(\tau)$ is an oscillatory integral which involves Bessel- like functions with different parameters, and the formula for $\mathbb{V}[\mathcal{N}_{J,\theta_0}]$ features quite complicated expressions using the first and second derivatives of this integral, with a global integration on top of this; so any analytical understanding of this formula seems out of reach ! But we can check numerically that it does attest to monochromatic fields having minimum variance.

We used Mathematica to evaluate variances of invariant GRFs, using the formulae in the theorem of Cramer and Leadbetter's. This needed some care: to evaluate $\mathbb{V}[\mathcal{N}_{J,\theta_0}]$, we had to perform numerical integration on an expression involving derivatives of the correlation function G, itself a parameter-dependent integral which cannot be reduced to simpler functions of the parameter. This kind of numerical evaluation is rather delicate to perform precisely, especially if there are oscillations in the integral as is the case here − the result can then be very highly dependent on the sampling strategy − and if there are multiple operations to be performed on the outputs of these integrals − the calculations of derivatives and second derivatives of the numerically- evaluated G, and the multiple divisions, might propagate the errors quite erratically.

In order to keep the numerical errors from masking the "exact" effect of thickening the spectrum, we forced the software to optimize its calculation strategy (adaptive Monte-Carlo integration), detecting oscillations in the integrand and adapting the sampling requirements, and we extended evaluation time beyond the usual limits (by dropping the in- built restrictions on the recursion depths). When the difference between successive evaluations was tamed, this yielded the variance curve displayed on fig. 4.

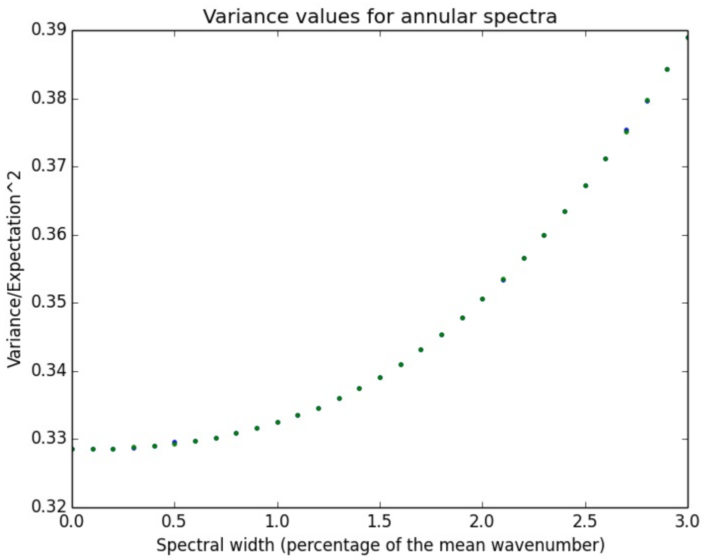


Figure 4: **Variance is a decreasing function of spectral thinness.** This is a plot of the variance of the random variable recording the number of times a given orientation is present on a straight line segment of fixed length. We considered here Invariant Gaussian Random Fields with uniform power spectra, and plotted the variance as a function of spectral width. For each percentage of the mean wavenumber, we displayed two outputs to give an idea of the attained precision. A low value for variance, here expressed in the unit given by the square of the expectation, corresponds to a field whose hypercolumns have relatively constant size across the resulting orientation map; the results displayed here show that a very regular hypercolumnar organization is quite compatible with stochastic modelling, and is a direct consequence of the spectral thinness condition found in models. Moreover, the horizontal slope at zero shows that as regards the global properties of quasiperiodic maps, there is very little difference between a theoretically ideal monochromaticity and a more realistic (and more model- independent) spectral thinness.

Note that the drawn variances correspond to fields with very slightly different spacings $\Lambda_{sq.}$; however, it is easy to check numerically that for every spectrum considered here, the variance of a monochromatic field with wavelength $\Lambda_{sq.}$ is inferior to the variance drawn on fig. 4.

Numerical evaluations also show that at a fixed spectral width, using few circles to build the field (i.e. introducing several characteristic wavelengths in the map) leads to a higher variance than simulating a uniform spectral distribution. To see this, we first evaluated $\mathbb{V}[\mathcal{N}_{J,\theta_0}]$ for an invariant GRF whose spectrum gathered three circles of radii R_{inf}, R_{sup} and $R_{mean} = 10.95$ in the fixed arbitrary unit, then spanned the interval between R_{inf} and R_{sup} with more and more circles, using spectra with 2N+1 circles of radii $R_{mean} + 0.95\frac{i}{N}$. We observed $\mathbb{V}[\mathcal{N}_{J,\theta_0}]$ to decrease with N in that case, and the existence of a limit value. From Riemann's definition of the integral, we see that this value is that which corresponds to a spectrum uniformly distributed in the annulus delimited by R_{inf} and R_{sup}. To keep the evaluation time reasonable (it is roughly quadratic in N), we kept the value $N = 18$ for the evaluations whose results are shown on fig. 3, and which are close to the observed limit values. We should also add here that we observed higher values for variance when using smooth spectra with several dominant wavelengths.

This is another argument for monochromaticity yielding minimum variance. Since the space of possible spectra with a fixed support is infinite-dimensional, our numerical experiments cannot explore it all. But we feel justified in stating the following numerical results on quasiperiodicity in orientation maps sampled from invariant gaussian fields:

Result 3 :

(i) For uniform spectra, variance increases with the width of the supporting annulus.

(ii) For a given spectral width, dominance of a single wavelength seems to yield minimum variance. Introducing more than one critical wavelength in the spectrum systematically increases nonuniformity in the typical size of hypercolumns.

Result 3 proves that sharp dominance of a single wavelength is the best way to obtain minimum variance. What is more, the horizontal slope at zero in fig. 3 means that fields which are close to monochromatic have much the same quasiperiodicity properties as monochromatic invariant fields. This is quite welcome in view of Schnabel's results: of course we cannot expect actual monochromaticity in real OPMs, but clear dominance of a wavelength is much more reasonable biologically. A more theoretical benefit is the flexibility of invariant GRFs for modelling: a model-adapted precise formula for the power spectrum may be inserted without damage to the global, robust resemblance between the predicted OPMs and real maps [11].

These observations reinforce the hypothesis that our three informational principles (randomness structure, invariance, spectral thinness) are sufficient to reproduce quantitative observable features of real maps, though as we saw, using an invariant GRF with the most realistic spectrum does not necessarily yield a more realistic result than using a monochromatic GRF, and leads to incompatibilities with the observed mature maps.. This form of universality is certainly welcome: individual maps in different animals, from different species (with different developmental scenarii) necessarily have different spectra, but general organizing principles *can* be enough to explain even quantitative observed properties.

3 Discussion

In this short note we recalled that simple hypotheses on randomness, invariance and spectral width of model orientation maps reproduce important geometrical features of real maps. Though it should not be forgotten that we worked in a simplified mathematical framework which reproduces only some aspects of the common design and whose dissemblance with real maps can be established rigorously as we recalled, we feel two new points deserve special attention: first, we showed that in the simplified setting of Gaussian Random Fields, the best mathematical quantity for explaining the local quasiperiod is the quadratic mean wavenumber rather than the mean wavenumber, and pointed out that a pinwheel density of π, when scaled with respect to this intrinstic column spacing, is a signature of Euclidean symmetry and not of Euclidean symmetry plus spectral thinness ; second, we established (through numerical analysis of an exact formula) that the variability of local quasiperiods is minimized when the standard deviation of the spectral wavelength tends to zero.

Our analysis shows that at least in the setting of Gaussian fields, realistically large spectra are compatible with a low variance; we suggest that a low variance for column spacing might be observed in real data, and perharps also a low variance for the number of pinwheels in an area Λ_{exp}^2. Spectral thinness is usually attributed to biological hardware in the cortex (like pre- sight propagation wavelengths in the retina or thalamus [37, 38]); this turns out to be compatible with some forme of optimality in information processing.

It would also be very interesting to compare the variance of column spacings in real maps (in units of the spacing evaluated by averaging local periods) with the smallest possible value for GRFs, observed in this paper for monochromatic fields (see figure 4); if a lower value for variance in real maps than in monochromatic Gaussian fields is found, it would mean that cortical circuitry refinment, featuring long- range interactions, brings mature maps closer to a geometrical homogeneity of hypercolumns. This would also throw some light on the fact that as development proceeds and the probability distribution of the field turns away from that of a GRF, driven by activity- dependent shaping, the column spacing obtained by averaging local periods seems to come closer to the wavelength associated to the mean or peak wavenumber (see [12], supplementary material, p. 5) than it is in GRFs. It is then remarkable that development should maintain the value of π for pinwheel density when scaled with respect to the current value of column spacing, keeping Result 2 valid over time (of course the density seems to move if one does not change the definition of column spacing over time, but the best- suited quantity for measuring column spacing seems to change). Perhaps this also has a benefit for areas receiving inputs from V1, keeping their tuning with the pinwheel subsystem (which seems to have an independent interest for information processing, see [45]) stable.

Acknowledgments

My doctoral advisor Daniel Bennequin supplied many suggestions, constant support, and very welcome help in writing this note: I owe him many warm thanks. It is a pleasure to thank Fred Wolf for useful discussion; I am also grateful to Christine Cachot for her help in setting up numerical experiments. The contents of this paper benefited greatly from the remarks of two anonymous referees: let me also thank them here.

Bibliography

[1] Hubel, D. H., & Wiesel, T. N. *Receptive fields, binocular interaction and functional architecture in the cat's visual cortex.* The Journal of physiology 1962, **160(1)**, 106.

[2] Hubel, D. H., & Wiesel, T. N. *Receptive fields of single neurones in the cat's striate cortex.* The Journal of physiology 1959 **148(3)** :574- 591.

[3] Hubel, D. H., & Wiesel, T. N. *Receptive fields and functional architecture of monkey striate cortex.* The Journal of physiology 1968, **195(1)**, 215-243.

[4] Van Hooser S D, Heimel J A, Nelson S B. *Functional cell classes and functional architecture in the early visual system of a highly visual rodent.* Progress in brain research 2005, **149** :127- 145.

[5] Bonhoeffer T, Grinvald A. *The layout of iso-orientation domains in area 18 of cat visual cortex: optical imaging reveals a pinwheel-like organization* J. Neurosci 1996. **13**, 4157

[6] Bonhoeffer and Grinvald A. *Iso-orientation domains in cat visual cortex are arranged in pinwheel-like patterns.* Nature 1991, **353(6343)** :429- 431.

[7] Ohki, K., Chung, S., Kara, P., Hubener, M., Bonhoeffer, T., & Reid, R. C. *Highly ordered arrangement of single neurons in orientation pinwheels.* Nature 2006, **442(7105)**, 925-928.

[8] Paik SB, Ringach D. *Retinal origin of orientation maps in visual cortex.* Nature Neuroscience 2011 **14**: 919- 925

[9] Nauhaus I, Nielsen K, Disney A, Callaway E. *Orthogonal micro-organization of orientation and spatial frequency in primate primary visual cortex.* Nature Neuroscience 2012 **15**, 1683- 1690

[10] Swindale N, Shoham D, Grinvald A, Bonhoeffer T, Hubener M. *Visual cortex maps are optimized for uniform coverage.* Nature Neuroscience 2000 **3**, 822 - 826.

[11] Kaschube, M. *Neural maps versus salt-and-pepper organization in visual cortex.* Current opinion in neurobiology 2014, **24**: 95-102.

[12] Kaschube, M., Schnabel, M., Lowel, S., Coppola, D. M., White, L. E., & Wolf, F: *Universality in the evolution of orientation columns in the visual cortex.* Science 2010, **330(6007)**:1113-1116.

[13] Nauhaus I, Nielsen K J. *Building maps from maps in primary visual cortex.* Current opinion in Neurobiology 2014 **24**:1- 6.

[14] Swindale, N.V. *A model for the formation of orientation columns.* Proc. Royal Society of London. Series B: Biological Sciences 1982 **215 (1199)**, 211- 230.

[15] Yu H, Farley B J, Jin D Z, & Sur M. *The coordinated mapping of visual space and response features in visual cortex.* Neuron 1995, **47(2)**:267- 280.

[16] Reichl L, Heide D, Lowel S, Crowley JC, Kaschube M, et al. *Coordinated Optimization of Visual Cortical Maps (I) Symmetry-based Analysis.* PLoS Comput Biol 2012 **8(11)**

[17] Petitot, J. *Neurogéométrie de la vision: modeles mathematiques et physiques des architectures fonctionnelles.* Paris: Editions Ecole Polytechnique; 2008.

[18] Swindale, N. V. *The development of topography in the visual cortex: A review of models.* Network: Computation in neural systems 1996, **7(2)**, 161-247.

[19] Chalupa, L. M., & Werner, J. S. (Eds.). (2004). The visual neurosciences (Vol. 1). Cambridge, MA: MIT press.

[20] Wolf, F., & Geisel, T.: *Spontaneous pinwheel annihilation during visual development.* Nature 1998, **395(6697)**: 73-78.

[21] Kaschube, M., Wolf, F., Geisel, T., & Lowel, S.: *Genetic influence on quantitative features of neocortical architecture.*. The Journal of Neuroscience 2002, **22(16)**:7206-7217.

[22] Wolf, F., & Geisel, T.: *Universality in visual cortical pattern formation.* . Journal of Physiology-Paris 2003, **97(2)**: 253-264.

[23] Kaschube, M., Schnabel, M., & Wolf, F. : *Self-organization and the selection of pinwheel density in visual cortical development.* New Journal of Physics 2008, **10(1)**, 015009.

[24] Abrahamsen, P. *A review of Gaussian random fields and correlation functions.* Norsk Regnesentral, 1995.

[25] Adler, R. J., & Taylor, J. E. *Random fields and geometry* Berlin: Springer; 2009.

[26] Azais J M, Wschebor M, *Level Sets and extrema of random processes and fields.* New York: John Wiley and Sons; 2009.

[27] Nauhaus I, Busse L, Carandini M, Ringach D. *Stimulus contrast modulates functional connectivity in visual cortex.* Nature neuroscience 2008 **12**(1) :70- 76.

[28] Schnabel M: *A Symmetry of the Visual World in the Architecture of the Visual Cortex.* PhD thesis. University of Goettingen; 2008.

[29] Yaglom, A. M. *Second-order homogeneous random fields.* In Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, Volume 2: Contributions to Probability Theory. The Regents of the University of California, 1961.

[30] Niebur, E., Worgotter, F., *Design principles of columnar organization in visual cortex.* Neural Computation 1994 **6 (4)**, 602- 614.

[31] Berry, M. V.; Dennis, M. R. *Phase singularities in isotropic random waves.* Royal Society of London Proceedings Series A 2000, **456(2001)**: 2059

[32] Azais, J. M., Leon, J. R., Wschebor, M. *Rice formulae and Gaussian waves.* Bernoulli 2011 **17(1)**, 170-193.

[33] Cramer, H. and Leadbetter, M.R. *Stationary and related stochastic processes. Sample function properties and their applications.* New York: John Wiley and Sons, 1967. Reprint: Dover books, 2004.

[34] Miller, K.D. $\pi =$ *Visual Cortex.* Science 2010 **330**:1059-1060.

[35] Bosking W H, Zhang Y, Schofield B, and Fitzpatrick D. *Orientation selectivity and the arrangement of horizontal connections in tree shrew striate cortex.* J. Neurosci. 1997, **17(6)**:2112- 2127.

[36] Barbieri, D., Citti, G., Sanguinetti, G., & Sarti, A. *An uncertainty principle underlying the functional architecture of V1.* Journal of Physiology-Paris 2012, **106(5)**, 183-193.

[37] Ernst, U. A., Pawelzik, K. R., Sahar-Pikielny, C., & Tsodyks, M. V. *Intracortical origin of visual maps.* Nature neuroscience 2001, **4(4)**, 431-436.

[38] Maffei, L., & Galli-Resta, L. *Correlation in the discharges of neighboring rat retinal ganglion cells during prenatal life.* Proceedings of the National Academy of Sciences 1990, **87(7)**, 2861-2864.

[39] Wolf F., *Symmetry, Multistability, and Long-Range Interactions in Brain Development.* PRL 2005, **95**: 208701

[40] Reichl L, Heide D, Lowel S, Crowley JC, Kaschube M, et al. *Coordinated Optimization of Visual Cortical Maps (II) Numerical studies.* PLoS Comput Biol 2012 **8(11)**

[41] Koulakov A, Chklovskii D, *Orientation preference patterns in mammalian visual cortex: a wire length minimization approach.* Neuron 2001 **29(2)**, pp. 519-27.

[42] Bohr, H. A. *Almost periodic functions.* New York: Chelsea Publishing Company, 1947.

[43] Harris, C. M., & Wolpert, D. M. *Signal-dependent noise determines motor planning.* Nature 1998, **394(6695)**, 780-784.

[44] Zhang, K., & Sejnowski, T. J. *Neuronal tuning: To sharpen or broaden ?* Neural Computation 1999, **11(1)**, 75-84.

[45] Dragoi V1, Rivadulla C, & Sur M. *Foci of orientation plasticity in visual cortex.* Nature 2001 **411(6833)**: 80-86.

Additional documentation

Sampling from monochromatic invariant random fields

In the main text, we defined monochromatic invariant Gaussian random fields through their correlation functions, and we studied the difference between monochromatic invariant fields and general invariant gaussian fields. On fig. 2 we displayed an OPM sampled from a monochromatic invariant gaussian field, but we did not say how the drawn object was built from its correlation function. We provide some details in this subsection.
Recall that the covariance function of a monochromatic invariant gaussian random field with correlation wavelength Λ is provided by the inverse Fourier transform of the Dirac distribution on a circle, that is,

$$\mathbb{E}\left[\mathbf{z}(x)\mathbf{z}(y)^\star\right] = \Gamma(x-y) = \int_{\mathbb{S}^1} e^{i\ R\ u\cdot(x-y)}du$$

with $R = \frac{2\pi}{\Lambda}$. Now, Γ satisfies the Helmholtz equation $\Delta\Gamma = -R^2\Gamma$; from this we can easily deduce that

$$\mathbb{E}\left[\left|\Delta\mathbf{z} + R^2\mathbf{z}\right|^2\right]$$

is identically zero. This means that any (strictly speaking, almost any) orientation map drawn from $\mathbf{z}$ satiisfies itself the Helmholz equation; thus OPMs drawn from $\mathbf{z}$ are superpositions of plane waves with wavenumber R and various propagation directions.
Thus, we know that there is a random Gaussian measure $d\mathbb{Z}$ on the circle which allows for describing $\mathbf{z}$ as a stochastic integral :

$$\mathbf{z}(x) = \int_{\mathbb{S}^1} e^{iR\ u\cdot x} d\mathbb{Z}(u).$$

Now, from the Gaussian nature of $\mathbf{z}$ and the Euclidean invariance condition, we have a simple way to describe $\mathbb{Z}$, which we used for actual computations: if $(\zeta_k)_{k\in\mathbb{N}^\star}$ is a sequence of independant standard Gaussian complex random variables, and if $u_1, ...u_n$ are the complex numbers coding for the vertices of a regular n- gon inscribed in the unit circle, then

$$\mathbf{z}_n : x \mapsto \frac{1}{n}\sum_{i=1}^{n} \zeta_i\ e^{iR\ u_i\cdot x}$$

is a Gaussian random field. As n grows to infinity, we get random fields which are closer and closer to being a monochromatic invariant gaussian random field, and our field $\mathbf{z}$ is but the limiting field.

Kac-Rice formula

We derived Result 1 from the classical Kac- Rice formula, and the theorem from which we obtained Result 2 can be obtained from a suitable generalization to plane random fields (see refs [26, 32] and [25, 22, 31] in the main text). Here we give the precise theorem we used in the derivation of Result 1. This formula was obtained as early as 1944, though the road to a complete proof later proved sinous; the initial motivation on Rice's side was the study of noise in communication channels, which can be thought of as random functions of time. For modelling noise it is then reasonable to introduce Gaussian random fields defined on the real line, and if the properties of the communication channel do not change over time, to assume further that they are stationary. Rice discovered that there is a very simple formula for the mean number of times this kind of field crosses a given "noise level"; this is the

Classical Kac-Rice formula : consider a stationary Gaussian Random Field Φ defined on the real line, with smooth trajectories; choose a real number u, and consider an interval I of length ℓ on the real line. Write $\mathcal{N}_{u,I}$ for the random variable recording the number of points x on I where $\Phi(x) = u$; then

$$\mathbb{E}\left[\mathcal{N}_{u,I}\right] = \ell \cdot \frac{e^{-u^2/2}\sqrt{\lambda}}{\pi}$$

where $\lambda = \mathbb{E}\left[\Phi'(0)^2\right]$ is the second spectral moment of the field.

For the proof of this old formula, as well as a presentation of all the features of GRFs underlying our main text, see ref. [25, 33]. For the proof of the theorem we used for Result 2, which is much more recent, we refer to refs. [22] and [32].
Just after Result 1, we mentioned that the comparison between Λ_{mean} and Λ_{sq} is an immediate consequence of Jensen's inequality, so let us give its statement here. We start with a continuous probability distribution $\mathbb{P}$ on the real line. Now whenever φ is a *convex*, real- valued function on $\mathbb{R}$, Jensen's inequality is the fact that for each measurable function f,

$$\varphi\left(\int_{\mathbb{R}} f(x)d\mathbb{P}(x)\right) \leq \int_{\mathbb{R}} \varphi\left(f(x)\right)\mathbb{P}(x).$$

(compare the triangle inequality). In the main text, we used it when $\mathbb{P}$ is the power spectrum distribution of a random field, f is the identity function of $\mathbb{R}$, and φ is $x \mapsto x^2$.

Chapter 3

Orientation maps in V1 and non-Euclidean geometry

Contents

To Jack Cowan, on the occasion of his 80th birthday.

Abstract

In the primary visual cortex, the processing of information uses the distribution of orientations in the visual input: neurons react to some orientations in the stimulus more than to others. In many species, orientation preference is mapped in a remarkable way on the cortical surface, and this organization of the neural population seems to be important for visual processing. Now, existing models for the geometry and development of orientation preference maps in higher mammals make a crucial use of symmetry considerations. In this paper, we consider probabilistic models for V1 maps from the point of view of Group theory; we focus on Gaussian random fields with symmetry properties and review the probabilistic arguments that allow to estimate pinwheel densities and predict the observed value of π. Then, in order to test the relevance of general symmetry arguments and to introduce methods which could be of use in modelling curved regions, we reconsider this model in the light of group representation theory, the canonical mathematics of symmetry. We show that through the Plancherel decomposition of the space of complex-valued maps on the Euclidean plane, each infinite-dimensional irreducible unitary representation of the special Euclidean group yields a unique V1-like map, and we use representation theory as a symmetry-based toolbox to build orientation maps adapted to the most famous non-Euclidean geometries, *viz.* spherical and hyperbolic geometry. We find that most of the dominant traits of V1 maps are preserved in these; we also study the link between symmetry and the statistics of singularities in orientation maps, and show what the striking quantitative characteristics observed in animals become in our curved models.

1 Introduction

In the primary visual cortex, neurons are sensitive to selected features of the visual input: each cell analyzes the properties of a small window in the visual field, its response depends on the local orientations and spatial frequencies in the visual scene [11, 12], on velocities or time frequencies [57, 58], it is subject to ocular dominance[11], etc. These receptive profiles are distributed among the neurons of area V1, and in many species they are distributed in a remarkably orderly way [21, 20, 32]. For several of these characteristics (position, orientation), the layout of feature preferences is two-dimensional in nature: neurons form so-called *microcolumns* orthogonal to the cortical surface, in which the preferred simulus orientation or position does not change [11]; across the cortical surface, however, the two-dimensional pattern of receptive profiles is richly organized [11, 21, 32, 15].

Amongst all feature maps in V1, it seems that the *orientation map* has a special part to play. Its beautiful geometrical properties (see fig. 1) have prompted many experimental and theoretical studies (see [1, 32, 21, 3, 10]); the orientation map seems to be closely tied to the horizontal wiring (the layout of connectivities between microcolumns) of V1 [32], its geometry is correlated to that of all the other feature maps [16], and while the geometrical properties of other feature maps vary much across species, those of orientation maps are remarkably similar [1].

It is thus tempting to attribute a high perceptual significance to the geometry of orientation maps, but is a long-standing mystery that V1 should develop this way: there are species in which no orientation map is present, most notably rodents [14, 27], though some of them, like squirrels, have fine vision [26]; on the other hand, it is a fact that orientation

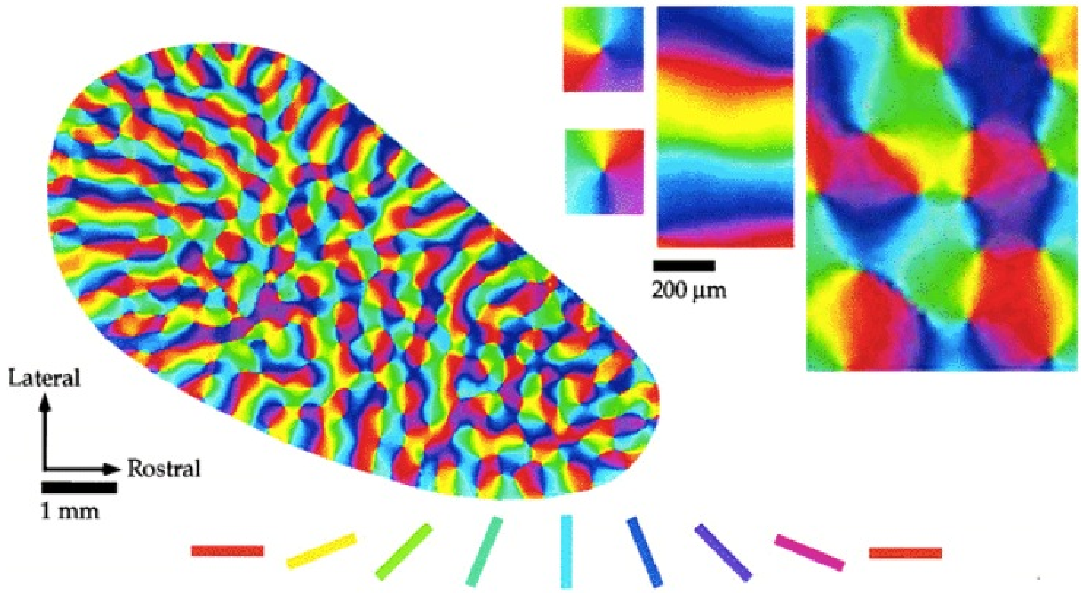


Figure 1: (figure modified from Bosking et al [32]). An Orientation Preference Map observed in the visual cortex of a tree shrew. The experimental procedure leading to this map is recalled in the main text. See also Swindale [17]. On the upper right corner, details at singular points (pinwheels) or regular points are shown.

maps are to be found in distantly related species whose common ancestor likely did not exhibit regular maps. This has led to an intense (and ongoing) debate on the functional advantage of these ordered maps for perception, on the conditions in which such maps develop, and on the part self-organization has to play in the individual (ontogenetic) development of V1-like geometries [1, 27].

Our concern here is not with these general issues, but on geometrical principles that underlie some elements of the debate. We focus on models which have been quite successful in predicting precise quantitative properties of V1 maps from a restricted number of principles.

Our results will be based on methods set forth by Wolf, Geisel and others while discussing development models. They have shown that the properties of mature maps in a large region of V1 (that which is most easily accessible to optical imaging) are well reproduced by treating the mature map as a sample from a random variable with values in the set of possible orientation maps, and by imposing symmetry conditions on this random variable (see the Methods section).

A remarkable chain of observations by Kaschube et al. [1, 2] has shown that there are universal statistical regularities in V1 orientation maps, including an intriguing mean value of π for their density of topological defects (with respect to their typical length of quasiperiodicity; see fig. 1 and the Methods section). Wolf, Geisel and others [3, 5, 6] give a theoretical basis for understanding this; one of its salient features is the use of Euclidean symmetry.

In this discussion, the cortical surface is treated as a full Euclidean plane. Then conditions of homogeneity and isotropy of the cortical surface are enforced by asking for the probability distribution of the mentioned random variable to be invariant under translations and rotations of this plane. This is a condition of invariance under the action of the Euclidean group of rigid plane motions.

There are several reasons for wondering why the cortical surface should be treated as a Euclidean plane, and not as a curved surface like the ones supporting non-Euclidean geometries.

The underlying assumptions are not explicitly discussed in the literature. For instance, rigid motions can be considered in
– the geometry of the visual field,
– the geometry of the cortical surface, that of the actual biological tissue,
– or an intermediate functional geometry (e.g. treating motions of solid objects against a fixed background).

It is true that the part of V1 which is accessible to optical imaging is mostly flat, and that we may imagine an affine visual field to be flat as well.

Above all we feel that these three “planes" should be carefully distinguished, and that using Euclidean geometry simultaneously at all levels is not without significance.

A first, casual remark is that the way the (spherical) retina records the visual field uses its projective properties; it is on a rather functional level that we wan think of “the" affine visual field related to it by central projection from the retina (eye movements are amazingly well-adapted to this reconstruction: see [33] for a discussion of motor computation in the Listing plane).

But more importantly, for almost all (if not all) animals which have been investigated, the correspondence between the accessible cortical region and the visual field (the retinotopic map) strongly departs from a central projection: it is logarithmic in nature, with a large magnification factor. For instance, even for the Tree Shrew which is known to have cortical V1 mostly flat, the observed region does not correspond to the center of the retina and the representation of the central field covers the major part of V1. A consequence is that Euclidean plane motions on the cortical surface and rigid motions in the visual field are very different. This is even more strikingly true for cats, primates and humans, whose calcarine sulcus has a more intricate 3D structure [23]. With this in mind, it seems very striking that the functional architecture of V1 should rely on a structure, the “association field" (see [10], chapter 4) and its condition of “coaxial alignment" of orientation preferences, that simultaneously uses Euclidean geometry at several of these levels. It is also very interesting to note the successful use of shift-twist symmetry (see section 3.4.2), a geometrical transformation which relates rotations on the cortical plane and rotations in the visual plane, in the study of hallucinatory patterns with contours [38] and of fine geometrical properties of V1 maps [8].

Thus when discussing plane motions, we feel that one should carefully keep track of the level (anatomical, functional, "external") to which they refer. On the other hand, it is quite clear in Wolf and Geisel’s development models that the Euclidean invariance conditions are imposed at the cortical level, independently of the retinotopic map [8, 1].

Is then using flat *Euclidean* geometry at the cortical level indispensable ? A closer look at the literature reveals that, when it appears, Euclidean geometry is endorsed only as a way to enforce conditions of *homogeneity and isotropy* on the two-dimensional surface of cortical V1. This makes it reasonable to look at the conditions of homogeneity and isotropy in non-Euclidean cases.

Now, these two notions are not at all incompatible with curvature; they are central in studying two-dimensional geometries with nonzero curvature, discovered and made famous by Gauss, Bolyai, Lobatchevski, Riemann and others. Extending the notions from geometry, analysis and probability to these spaces has been a source of great mathematical achievements in the late 19th and throughout the 20th century (Lie, Cartan, Weyl, Harish-Chandra, Yaglom). The central concept is that of transformation group, and the corresponding mathematical tools are those of noncommutative harmonic analysis, grounded on Lie group representations. In fact, Wolf and Geisel's ingredients precisely match the basic objects of invariant harmonic analysis.

Our aim in this paper is to use these tools to define natural V1-like patterns on non-Euclidean spaces. Because symmetry considerations are central to the whole discussion, we need our non-Euclidean spaces to admit enough symmetries for the conditions of homogeneity and isotropy to make sense, and we thus consider the two-dimensional symmetric spaces. Aside from the Euclidean plane there are but two continous families of models for such spaces, isomorphic to the sphere and the hyperbolic plane, so these two spaces will be the non-Euclidean settings for our constructions.

The success of Euclidean-symmetry-based arguments for describing flat parts of V1 makes it quite natural, from a neural point of view, to wonder whether in curved regions of V1, the layout of orientation preferences develops according to the same principles, and what could be the importance of the metric induced by cortical folding or of "coordinates" which would be induced by flattening the surface (and with respect to which the notion of curvature loses its meaning). It is a matter of current debate whether the three-dimensional structure induced by cortical folding has functional benefits; present understanding seems to be that its structure is the result of anatomical constraints (like the tension along cortico-cortical connections, or the repartition of blood flow, see [72]), but several hypotheses have been put forward to assess its functional meaning (see for instance [70, 72]). In a study trying to assess the importance of cortical folding for orientation maps it would be natural of course to consider variable curvature, but it is difficult to see how symmetry arguments could generalize and even make sense, whereas in regions having large (local) symmetry groups we shall see that it is very natural to adapt the successful arguments for flat V1 after a suitable interpretation of the latter. As we shall point out in the upcoming Discussion, there might also be benefits (in terms of information processing) in having symmetry groups as large as possible in rather extended regions.

Here is an outline of the paper. In the Methods section, we first proceed to describe some aspects of Wolf and Geisel's models with the words of representation theory; in this situation the relevant group is the Euclidean group of rigid plane motions. We introduce the probabilistic setting to be used in this paper, that of Gaussian Random Fields, in subsection 2.1, and discuss the crucial Euclidean symmetry arguments in section 2.2. We bring group theory into the picture in section 2.3, and irreducible representations in subsection 2.4.

To pass over to non-Euclidean geometries, we then examine what happens if the Euclidean group is replaced by the isometry groups of other symmetric spaces; we thus define "orientation maps" on surfaces of negative or positive curvature. For symmetric spaces the curvature is a numerical constant, and after a renormalization the two-dimensional symmetric spaces turn out to be isomorphic with the Euclidean plane, the hyperbolic plane or the round sphere. We begin the Results section with the hyperbolic, negatively curved setting rather than the spherical, positively-curved one, because there are closer links with flat harmonic analysis in that case. After introducing our orientation-preference-like maps on these spaces, we emphasize the important part symmetry plays in the existence of the universal value for defect (pinwheel) densities in V1 maps by discussing the density of topological defects in non-Euclidean orientation maps.

As we shall see, in the Euclidean case, irreducible representations enter the picture through the existence of a dominant wavelength in the correlation spectrum; our recent paper in this journal [73] focuses on the role of this monochromaticity condition in getting a precise pinwheel density and quasiperiodicity. Although some of our results can find motivation from a few remarks in that paper, the present study is independent from [73].

Acknowledgments. I owe many warm thanks to my doctoral advisor Daniel Bennequin, who suggested this work and has been of much help in designing and writing this paper. I am grateful to the referees whose help was precious for making this paper more readable. I also thank Alessandro Sarti for discussions of Euclidean symmetry in V1, and Michel Duflo for his kind interest in the results of this paper.

2 Methods

2.1 Gaussian random fields

How was the map of fig. 1 obtained [32] ? High-contrast square wave gratings were presented to the animal, and optical imaging was used to measure the difference between the responses of neurons on the cortical surface upon translation of the visual input. From these data, a pattern emerges that attributes, given a stimulus orientation ϑ, a sensitivity $a_\vartheta(x)$ to every point x of the cortical surface (so a_ϑ is a positive-valued continous function on the cortical surface $\mathcal{X}$). If this is recorded for a number of directions $\vartheta_1, ...\vartheta_N$, and if the column beneath a point $x_0 \in \mathcal{X}$ of the cortical surface has orientation preference ϑ_j, then the polygon whose vertices are the points $a_{\vartheta_k}(x_0)e^{2i\vartheta_k}$ in $\mathbb{C}$ will be elongated in the direction $2\vartheta_j$, and the argument of the complex sum

$$\mathbf{z}_{exp}(x_0) = \sum_{k=1}^{N} a_{\vartheta_k}(x_0)e^{2i\vartheta_k}$$

will be approximately $2\vartheta_j$. The functions a_{ϑ_k} for a tree shrew V1 were obtained using optical imaging, and the map drawn on fig 1 is simply $x \mapsto \frac{1}{2}\arg \mathbf{z}_{exp}(x)$.

Let us add that if there is a pinwheel center at x_0, by definition [1] $\mathbf{z}_{exp}$ takes all values of the argument in a neighbourhood of x_0, so $\mathbf{z}_{exp}(x_0)$ must be zero. On the other hand, the modulus of $\mathbf{z}_{\mathrm{exp}}$ may loosely be interpreted as a measure of orientation selectivity: when orientation tuning at x_0 is poor, all of the $a_{\vartheta_k}(x_0)$ will be approximately the same, so x_0 will be close to a zero of $\mathbf{z}_{exp}$, while if orientation selectivity at x_0 is sharp, the numbers $a_{\vartheta_k}(x_0)$ for which ϑ_k is close to the preferred orientation will be much larger than the others, and the modulus of $\mathbf{z}$ will be rather high at x_0.

With this interpretation, we may discuss any complex-valued smooth function $\mathbf{z}$ on a surface X as if its argument were an orientation map, and its modulus were a measure of orientation selectivity. Orientation selectivity near pinwheel centers is being actively researched and debated, see [15, 65] and the references in [66], so interpreting the modulus of the vector sum z_{exp} in this way might be questioned, but this tradition dates back to 1982 [17].

1. Note that $\mathbf{z}_{exp}$ is automatically continous.

If mathematical models yielding plausible maps z are to be furnished, then certainly they should be compared to the multitude of maps observed in different individuals. Let us neglect, for a given species, the slight differences in cortical shape and assume that each test subject comes with a coordinate system on the surface of its $V1$, so that we may compare a given map from $\mathbb{R}^2$ to $\mathbb{C}$ to the orientation map observed in this individual.

We can then compare the different individual maps, leading to *map statistics*; if orientation maps are to be described mathematically, it seems fair to hope for a mathematical object that produces, rather than a single complex-valued function with the desired features, statistical ensembles of realistic-looking maps [3]. This approach might not be the best way to account for the finer properties of mature maps as experimentally observed, and it is certainly a rough approximation that needs to be confronted with the output of more biologically plausible development models. However, it does have the advantage of mathematical simplicity, and as we shall see, it is particularly well-suited to discussing the part symmetry arguments have to play in producing realistic maps.

So what we need is a *random field*, that is, a random variable with values in the set of smooth maps from $\mathbb{R}^2$ to $\mathbb{C}$. Since the set of smooth maps is infinite-dimensional, we cannot expect to find interesting “probability distributions" from closed formulae [28, 30]; but in the case of V1, the general theory of random fields and the available biological information make it possible to describe special fields whose “typical realizations" yield rather realistic maps [30, 5]. When we go over to non-Euclidean settings in this paper, we shall see that the mathematical description can be adapted to provide special random fields defined on non-Euclidean spaces; their typical realizations will yield V1-like maps adapted to the considered non-Euclidean geometry.

But let us now make our way towards the special fields on Euclidean space whose typical realizations look like orientation maps.

Measured statistical properties of real orientation maps include correlation functions [8]: it turns out that the structures of correlation measured in different individuals look very much alike. This is important : many discussions take the architecture of correlations to be essential to the horizontal wiring of V1, and to be at the heart of its perceptual function [10, 22]; it is also at the heart of striking results on the distribution of singularities in OPMs [3, 5, 1]. So using models that reproduce this correlation structure seems to be a good idea, and there is a way to associate special random fields to correlation structures:

Definition: A complex-valued random function $\mathbf{z}$ on a smooth manifold M (a collection $\mathbf{z}(x), x \in M$ of complex-valued random variables) is a *complex-valued centered gaussian random field* (GRF) if, for every integer n and every n-tuple $(x_1, ...x_n) \in M^n$, the $\mathbb{C}^n$-valued random variable $(\mathbf{z}(x_1), ...\mathbf{z}(x_n))$ is gaussian with zero mean. Its *correlation function* $C : M^2 \to \mathbb{C}$ is the (deterministic) map $(x, y) \mapsto \mathbb{E}\left[\mathbf{z}(x)\bar{\mathbf{z}}(y)\right]$.

Just as a gaussian probability distribution on $\mathbb{R}$ is available when a value for expectation and a value for variance are given (and is the "best bet", that is the minimum entropy distribution, given these data [52]), a continuous two-point correlation function $C : M^2 \mapsto \mathbb{C}$ (together with the zero-mean requirement in the defintion we use here) determines a unique GRF thanks to an existence theorem by Kolmogorov: see [60], Theorem 12.1.3,

[30], and [28], p.4. [2]

In what follows, we shall always require that $C : M^2 \to \mathbb{C}$ be smooth enough; in fact we will only meet fields with real-analytic correlation functions. Maps drawn from such fields are almost surely smooth, so there is no regularity problem ahead.

Before we add symmetry constraints on our gaussian fields, note that $C(x, x)$ is the variance of $\mathbf{z}(x)$; this depends on a choice of unit for measuring orientation selectivity. We shall proceed to a convenient one in the next subsection.

2.2 Euclidean symmetry in V1

Let us for the moment deal with the cortical surface as if it were an Euclidean plane $\mathbb{R}^2$. In a grown individual, different points on this plane correspond to neurons that usually do not have the same orientation preference, whose connectivity reaches out to different subsets of the cortex [10, 22, 38]; at some points we find sharp orientation tuning and under others (pinwheels) a less clear behaviour. In short, two different points on the cortical surface usually have different parts to play in the processing of visual information. But experimental evidence [32, 19] suggests very clearly that no particular point on this plane should have any distinguished part to play in the *general design* of the orientation map (e.g. be an organizational center for the development of the map, or have a systematic tendency to exhibit a particular orientation preference in the end).

These two facts are *not* incompatible: we may use this homogeneity condition as a constraint on the *ensemble properties* of the gaussian field we are trying to obtain realistic maps from. In other words, given a possible outcome $x \mapsto z(x)$ for $\mathbf{z}$, we may ask that

$$x \mapsto z(x + u) \quad \text{(where u is any vector in } \mathbb{R}^2\text{)}$$

and

$$x \mapsto z(Rx) \quad \text{(where R is any } 2 \times 2 \text{ rotation matrix)}$$

have the same occurrence probability as $x \mapsto z(x)$. Rotations and translations come together in the *Euclidean group* SE(2), which is the set of transformations of the plane that preserve Euclidean distance and the orientedness of bases [40, 43, 48]; an element g of this group is easily shown to be uniquely specified by a couple (R, u) where R is a rotation matrix and u a vector, and it is readily checked that elements $g_1 = (R_1, u_1)$ t $g_2 = (R_2, u_2)$ compose as $g_1 g_2 = (R_1 R_2, u_1 + R_1 u_2)$.

The above assumption is then that *the probability distribution of* $\mathbf{z}$ *is invariant under the action of E(2)* on the set of maps [5].

This implies that $C(x, y)$ depends only on $\|x - y\|$, and in the case of a Gaussian field this apparently weaker form of invariance is actually equivalent to the invariance of the full probability distribution. Let us write $\Gamma : \mathbb{R}^2 \to \mathbb{C}$ for the radial function such that $C(x, y) = \Gamma(x - y)$ for all x and y, and note that up to a global rescaling of the modulus used to measure orientation selectivity, we may (and will) assume $\Gamma(0) = 1$.

Further discussion of correlations may be conducted using Γ, and there is an important remark to be made here: the high-frequency components of its Fourier transform record *local* correlations, while low-frequency components in $\widehat{\Gamma}$ (the Fourier transform of Γ) point

2. There are some conditions for this, called positive-definiteness, but they are automatically satisfied by correlation functions obtained from experimental data

to long-range correlations. If $\mathbf{z}$ is to produce a quasi-periodic layout of orientation preferences with characteristic distance Λ, this seems to leave no room for systematic correlations at a much longer or much shorter distance than Λ. So it seems reasonable to expect that gaussian fields generating plausible maps have $\widehat{\Gamma}$ supported on the neighborhood of a circle with radius $\frac{2\pi}{\Lambda}$. Following Niebur and Worgotter, Wolf and Geisel and others, we note that *this further hypothesis on* $\widehat{\Gamma}$ *is all that is needed to generate realistic-looking maps.*

The simplest way to test this claim is to use what we shall call a *monochromatic invariant random field*, a field in which $\widehat{\Gamma}$ actually has support in a single circle, and consequently is the Dirac distribution on this circle [3] : $\Gamma(\vec{r}) = \displaystyle\int_{\mathbb{S}^1} e^{i\frac{2\pi}{\Lambda}\vec{u}\cdot\vec{r}} d\vec{u}$.

This determines a unique GRF $\mathbf{z}$ (see the appendix for details on how to construct it from Γ), so let us draw an orientation map from this $\mathbf{z}$: the result is shown on figure 2.

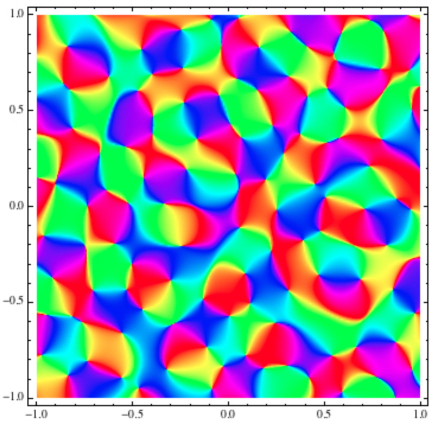


Figure 2: **Computer-generated map, sampled from a monochromatic field**. This figure shows an orientation map which we have drawn from a simulated Invariant Gaussian Random Field with circular power spectrum. This figure was generated using a superposition of 30 plane waves with frequency vectors at the vertices of a random polygon inscribed in a circle, and random Gaussian weights (see the appendix); what is plotted is the argument. In the unit of length displayed on the x-and y-axes, the wavelength is 1/3 here.

This looks realistic enough. Now, what is truly remarkable is that it is not only on a first, qualitative look that this map − which has been computer-generated from simple principles − has the right features: it also exhibits a pinwheel density of π, which Kaschube et al. have observed in real maps with 2% precision [1].

Indeed, if $\mathcal{N}_A$ is the random variable recording the number of pinwheels in an region A with area $|A|$, it may be shown (using a formula of Kac-Rice type, see [30, 31] for background on the Kac-Rice formula) that

$$\mathbb{E}\left\{\frac{\mathcal{N}_A}{|A|}\right\} = \frac{\pi}{\Lambda^2}.$$

This result appeared in physics [3, 24, 5] and is now supported by full mathematical rigor (see [25], chapter 6, for the general setting and [29], section 4 for the full proof); we shall use the same methods to derive non-Euclidean pinwheel densities in the Results sections.

3. Recall that $\widehat{\Gamma}$ is rotation-invariant

We should note here that as translation-invariant random fields of our type have ergodicity properties (see [30], section 6.5), it is quite reasonable to compare ensemble expectations for gaussian fields and pinwheel densities which, in experiments, are measured on individual orientation maps.

Of course the correlation spectra measured in real V1 maps are not concentrated on an infinitely thin annulus (for precise measurements, see Schnabel [8], p 103). But upon closer examination (see for instance [73]), one can see that maps sampled from invariant gaussian fields whose spectra are not quite monochromatic, but concentrated on thin annuli, do not only look the same as that of figure 2, but that many interesting quantitative properties (such as pinwheel density or a low variance for the spacing between iso-orientation domains) have vanishing first-order terms as functions of spectral thickness. In other words, it is reasonable to say that monochromatic invariant random fields provide as good a description for the layout of orientation preferences as invariant fields with more realistic spectra do (perhaps even better, see [73]). As we shall see presently, neglecting details in the power spectrum and going for maximum simplicity allows for a generalization that will lead us to pinwheel-like arrangments in non-Euclidean settings. We shall take this step now and start looking for non-Euclidean analogues of monochromatic fields.

But to sum up, let us insist that three hypotheses introduced in [3] gave map ensembles with realistic qualitative and quantitative properties:

(1) a *randomness structure*, that of a smooth gaussian field;

(2) an assumption of *euclidean invariance*;

(3) and a *monochromaticity, or near-to-monochromaticity condition* out of which quasi-periodicity in the map arose.

When we go over to non-Euclidean settings, these are the three properties that we shall look for. The first only needs the surface on which we draw orientation maps to be smooth. For analogues of the last two conditions in non-euclidean geometries we need group actions, of course, and a non-Euclidean notion of monochromatic random field. In the next two subsections, we shall describe the appropriate tool.

Before we embark on our program, let us note that in spite of the close resemblance between maps sampled from monochromatic Gaussian fields and real mature maps, there are notable differences. As we remarked above, real correlation spectra are not infinitely thin, and the precise measurements by Schnabel make it possible to give quantitative arguments for the difference between an invariant Gaussian field with the measured spectrum (see for instance a discussion in [73]). In the successful long-range interaction model of Wolf, Geisel, Kaschube and coworkers, Gaussian fields turn out to be a better description of the initial stage of cortical map development than they are of the mature stage. We have two reasons for sticking to Gaussian fields in this paper: the first is that they are ideally suited to discussing and generalizing the concepts crucial to producing realistic maps, and the second is that a non-Euclidean version of the long-range interaction model can easily be written down in the upcoming Discussion.

2.3 Klein geometries

What is a space M in which conditions (1) and (2) have a meaning ? Condition (1) says we should look for gaussian fields whose trajectories yield smooth maps, so M should be a smooth manifold. To generalize condition (2), we need a group of transformations acting on M, with respect to which the invariance condition is to be formulated. Felix Klein famously insisted that the geometry of a smooth manifold M on which there is a transitive group action is completely determined by a pair (G, K) in which G is a Lie group and K a closed subgroup of G. We shall recall here some aspects of Klein's view, focusing for the two-dimensional examples which we will use in the rest of this paper. This is famous and standard material; see the beautiful book by Sharpe [56], chapter 4.

First, let us examine the previous construction and note that every geometrical entity we met can be defined in terms of the Euclidean group $SE(2)$. write K for the subgroup of rotations around a given point, say o. If $g = (R, \vec{x})$ is any element of G, the conjugate subgroup $gKg^{-1} = \{(A, \vec{x} - A\vec{x}), A \in K\}$ is the set of rotations around $o + \vec{x}$. Now, the set $gK = \{(A, x), A \in K\}$ remembers x and only x, so we can recover the Euclidean plane by considering the family of all such cosets, that is, the set $G/K = \{gK, g \in G\}$.

Now when G is a general Lie group and K is a closed subgroup, the smooth manifold $M = G/K$ comes with a natural transitive G-action, and K is but the subset of transformations which do not move the point $\{K\}$ of M. This is summarized by saying that M is a G-homogeneous space.

With this in mind, we can rephrase our main objective in this paper: it is to show that some Klein pairs (G, K) allow for a construction of V1-like maps on the homogeneous space $M = G/K$, and a calculation of pinwheel densities in these V1-like maps. We shall keep M two-dimensional here, and stick to the three maximally symmetric spaces [44] – the Euclidean plane, the round sphere and the hyperbolic plane.

To recover the usual geometry of the round sphere $\mathbb{S}^2$ from a Klein pair, we need the group of rotations around the origin in $\mathbb{R}^3$, that is $G = SO(3) = \{A \in \mathcal{M}_3(\mathbb{R}) | {}^t AA = I_3 \text{ and } \det(A) = 1\}$, and the closed subgroup $K = \left\{ \begin{pmatrix} R & 0 \\ 0 & 1 \end{pmatrix}, R \in SO(2) \right\}$ – of course K is the group of rotations fixing $(0, 0, 1)$.

Let us now give some quick details on how the hyperbolic plane can be defined from a Klein pair. Here G is the group of linear transformations of $\mathbb{C}^2$ that have unit determinant and preserve the quadratic form $(z, z') \mapsto |z|^2 - |z'|^2$, that is,

$$G = SU(1,1) = \left\{ \begin{pmatrix} \alpha & \beta \\ \bar{\beta} & \bar{\alpha} \end{pmatrix} \mid \alpha, \beta \in \mathbb{C}, |\alpha|^2 - |\beta|^2 = 1 \right\}.$$

Elements of G operate on the complex plane $\mathbb{C}$ *via* conformal (but non-linear) transformations: any element $g = \begin{pmatrix} \alpha & \bar{\beta} \\ \beta & \bar{\alpha} \end{pmatrix}$ of G gives rise to a homography $z \mapsto g \cdot z := \frac{\alpha z + \beta}{\bar{\beta} z + \bar{\alpha}}$ of the complex plane. It is easy to see that the origin can be sent anywhere on the *(open) unit disk*, but nowhere outside. Now the subgroup K of transformations that leave 0 invariant is obviously $K = \left\{ \begin{pmatrix} e^{i\varphi} & 0 \\ 0 & e^{-i\varphi} \end{pmatrix}, \phi \in \mathbb{R} \right\}$; note that its elements induce the ordinary rotations of the unit disk. So the unit disk $\mathbb{D}$ in $\mathbb{C}$ comes with a Klein pair (G, K), and looking for a G-invariant metric on $\mathbb{D}$ famously produces the negatively-curved *Poincaré metric* (see [**?**, 40, 45] and the appendix for details). Recall that the formula for the square of the length of a vector tangent to $\mathbb{D}$ at (x, y) is summarized by

$$ds^2 = \frac{4}{(1-(x^2+y^2)^2)^2}(dx^2+dy^2).$$

We shall write $\eta(p)$ for the numerical function $p = (x, y) \mapsto \frac{4}{(1-(x^2+y^2)^2)^2}$, and η for the abstract G-invariant metric we have just defined. Of course group theory allows for a simple description of all its geodesics and of many other things geometrical, and we shall need those: to keep the size of the present section reasonable, we delay this description to section 2.1.

But now let us go forward to meet one of the many reasons why Klein's description of spherical and hyperbolic geometries, far from being a matter of aesthetics, shows our concrete tools of map engeeniering the way to non-Euclidean places.

2.4 Group representations and non-commutative harmonic analysis

2.4.a

Assume we are given one of the two non-Euclidean Klein pairs (G, K) above and we wish to build an orientation map with properties (1), (2) and (3) from section 2.2. Conditions (1) and (2) say we should use a smooth complex-valued gaussian random field that is invariant under the action of G. We shall come back to exploiting condition (2) in time. But condition (3) depends on classical Fourier analysis, which is based using plane waves and thus seems tied to $\mathbb{R}^n$.

Fortunately there *is* a completely group-theoretical description of classical Fourier analysis too: for details, we refer to the beautiful survey by Mackey [54]. One of its starting points is the fact that for functions defined on $\mathbb{R}^n$, the Fourier transform turns a global translation of the variable (that is, passage from a function f to the function $x \mapsto f(x-x_0)$) into multiplication by a universal (nonconstant) factor (the Fourier transform $\widehat{f}$ is turned into $k \mapsto e^{ikx_0}\widehat{f}(k)$). From this behaviour of the Fourier transform under the *action of the group of translations*, some of those properties in Fourier analysis which are wonderful for engineering – like the formula for the Fourier transform of a derivative – follow immediately.

For many groups, including $SO(3)$ and $SU(1,1)$ which we will use in this paper, there is a "generalized Fourier transform" which gives rise to analogues of the property we just emphazised, although it is technically more sophisticated than classical Fourier analysis. It is best suited to analyzing functions defined on spaces with a G-action, yielding concepts of "generalized frequencies" appropriate to the group G.

It will then come as no surprise that the vocabulary of noncommutative harmonic analysis is well-suited to describing the invariant Gaussian field model for orientation preference maps in V1, since the key features of this model rest on the action of $SE(2)$ on the function space of orientation maps. As soon as we give details, it will also be apparent that an analogue of the monochromaticity condition (3) can be formulated in terms of these "generalized frequencies".

Before we discuss its significance and its relevance to Euclidean (and non-Euclidean) orientation maps, we must set up the stage for harmonic analysis; so we beg our reader for a little mathematical patience until section 2.4.3 brings us back to orientation maps.

Let G be a Lie group. Representation theory starts with two definitions: a *unitary representation of* G is a continuous homomorphism, say T, from G to the group $\mathcal{U}(\mathcal{H})$ of linear isometries of a Hilbert space; we write $(\mathcal{H}, T)$ for it. This representation is *irreducible* when there is no $T(G)$-invariant closed subspace of $\mathcal{H}$ except $\{0\}$ and $\mathcal{H}$.

We need to give two essential examples, the second of which is crucial to the strategy of this paper:

1. if p is a vector in $\mathbb{R}^n$, define $T_p(x) = e^{ip\cdot x}$ for each $x \in G = \mathbb{R}^n$; this defines a continuous morphism from $G = \mathbb{R}^n$ to the unit circle $\mathbb{S}^1$ in $\mathbb{C}$; by identifying this unit circle with the set of rotations of the complex line $\mathbb{C}$, we may say that $(\mathbb{C}, T_p)$ is an irreducible, unitary representation of $\mathbb{R}^n$. In fact, every irreducible unitary representation of $\mathbb{R}^n$ reads $(\mathbb{C}, T_p)$ where p is a vector. Thus, the set of irreducible representations of the group of translations on the real line or of an n-dimensional vector space is nothing else than the set of "time" or "space" frequencies in the usual sense of the word. [4]
2. Suppose M is the real line, the Euclidean plane, the sphere or the hyperbolic plane, and G the corresponding isometry group. If f is a complex-valued function on M, define $L(g)f := x \mapsto f(g^{-1} \cdot x)$. Then for every $g \in G$, $L(g)$ defines a unitary operator acting in the Hilbert space $\mathbb{L}^2(M)$ (here integration is with respect to the measure determined by the metric we chose on M); so we get a canonical unitary representation $(\mathbb{L}^2(M), \mathcal{L})$ of G. It is very important to note that this representation is *reducible* in our four cases; we discuss its invariant subspaces in the next subsection.

A word of caution: our first example , although it is crucial to understanding how representation theory generalizes Fourier analysis, is much too simple to give an idea of what *irreducible* representations of nonabelian groups are like: for instance, the space $\mathcal{H}$ of an irreducible representation very often happens to be infinite-dimensional, and this will be crucial in our discussion of hyperbolic geometry.

2.4.b

Suppose M is the Euclidean plane, the hyperbolic plane or the sphere. We shall now give an outline of the *Plancherel decomposition of* $\mathbb{L}^2(M)$, which is crucial to our strategy for producing non-Euclidean orientation maps. This is standard material: for details, we refer to [45], chapter 0.

Let us consider the representation $\mathcal{L}$ of example 2. above, acting on $\mathcal{H} = \mathbb{L}^2(M)$. Since is not irreducible, we may write $\mathcal{H} = \mathcal{H}_a \oplus \mathcal{H}_b$ where $\mathcal{H}_a$ and $\mathcal{H}_b$ are mutually orthogonal, *stable* subspaces of $\mathcal{H}$ (note that for this to be so, they must be closed), and try to decompose $\mathcal{H}_a$ further. We may hope to come to a *decomposition into irreducibles*, and hope to eventually be able to write

$$\mathcal{H} = \bigoplus_{\gamma} \left(\bigoplus_{i=1}^{m_\gamma} \mathcal{H}_{\gamma,i} \right),$$

a direct sum of invariant, mutually orthogonal subspaces $\mathcal{H}_{\gamma,i}$ which inherit *irreducible* representations of G from $\mathcal{L}$, with $\mathcal{H}_{\gamma,i}$ equivalent [5] to $\mathcal{H}_{\gamma',i'}$ if and only if $\gamma = \gamma'$.

When G is the rotation group $SO(3)$ and M is the sphere, or more generally when G is compact, this is actually what happens, and it is a part of the Peter-Weyl theorem that the above direct sum decomposition holds. In the case of the sphere, all the m_γs will be

4. Likewise, if n is an integer, define $T(u) = u^n = e^{in\vartheta}$ for every element $u = e^{i\vartheta} \in \mathbb{S}^1$; the circle $\mathbb{S}^1$ is a group under complex multiplication and $(\mathbb{C}, T)$ provides an irreducible unitary representation of $\mathbb{S}^1$.

5. Two given irreducible unitary representations $(\mathcal{H}_1, T_1)$ and $(\mathcal{H}_2, T_2)$ are *equivalent* if there is a unitary map U from $\mathcal{H}_1$ to $\mathcal{H}_2$ such that $UT_1(g)$ and $T_2(g)U$ coincide for every $g \in G$. In example (a), it is very easy to check that there is no such unitary map intertwining T_p and T_q if $p \neq q$).

equal to one and we will describe the $\mathcal{H}_\gamma$ in section 3.2.1. But for the noncompact groups $SE(2)$ and $SU(1,1)$, the decomposition process turns out to degenerate.

A simpler example will help us understand the situation: consider the representation $\mathcal{L}$ of $\mathbb{R}$ on $\mathbb{L}^2(\mathbb{R})$ (example (b) above). Since a change of origin induces but a (nonconstant) phase shift in the Fourier transform, the subspace $\mathcal{F}_I$ of functions whose Fourier transform has support in interval I, is invariant by each of the $\mathcal{R}(x), x \in \mathbb{R}$. But now it is true also that, say $\mathcal{F}_{[0,1]} = \mathcal{F}_{[2,2.5]} \oplus \mathcal{F}_{[2.5,3]} = \mathcal{F}_{[2,2.25]} \oplus \mathcal{F}_{[2.25,2.5]} \oplus \mathcal{F}_{[2.5,2.75]} \oplus \mathcal{F}_{[2.75,3]}$, and so on. Since we can proceed to make the intervals smaller and smaller, we see that an irreducible subspace should be a one-dimensional space of functions which have only one nonzero Fourier coefficient, in other words, each member of the irreducible subspaces should be a plane wave... which is *not* a square-integrable function ! So in this case, there is *no* invariant subspace of $\mathbb{L}^2(\mathbb{R})$ that inherits an irreducible representation from $\mathcal{R}$, and it is only by getting out of the original Hilbert space that we can identify irreducible “consituents" for $\mathbb{L}^2(\mathbb{R})$.

When M is the Euclidean plane or the hyperbolic plane, this is what will happen: starting from $\mathbb{L}^2(M)$, we shall meet spaces $\mathcal{E}_\omega$ of *smooth* (and *a priori* not square-integrable) functions which

– are invariant under the canonical operators $\mathcal{L}(g)$, $g \in G$,
– carry irreducible unitary representations of G,
– and together give rise to the following version of the Plancherel formula: for each $f \in \mathbb{L}^2(M)$ and for almost every x in M,

$$f(x) = \int_{\omega \in \mathcal{F}} f_\omega(x) d\Pi(\omega)$$

where $\mathcal{F}$ is some set of equivalence classes of representations of G (the "frequencies"), Π is a measure on $\mathcal{F}$ (the "power spectrum"), and for each $\omega \in \mathcal{F}$, f_ω is a member of $\mathcal{E}_\omega$ (a smooth function, then).

Recall that our aim in introducing noncommutative harmonic analysis is to find an analogue of the monochromaticity condition (3), section 2.2, in spherical and hyperbolic geometry. As we shall see presently, the situation in the Euclidean plane makes it reasonable to call an element of $\mathcal{E}_\omega$ or $\mathcal{H}_\gamma$ a *monochromatic map.* Belonging to one of the $\mathcal{E}_\omega$, resp. one of the $\mathcal{H}_\gamma$, will be our non-Euclidean analogue of the monochromaticity condition (3) in hyperbolic geometry, resp. spherical geometry. We shall see that a gaussian random field providing orientation-preference-like maps may be associated to each of these spaces of monochromatic maps, and that it yields quasi-periodic tilings of M with Euclidean-like pinwheel structures.

2.4.c

Let us now proceed to relate the Plancherel decomposition of $\mathbb{L}^2(\mathbb{R}^2)$ to the monochromaticity condition (3) in section 2.2. In the notations of section 2.2, (3) means that the correlation function Γ of a monochromatic field should have its support on the circle of radius $\frac{2\pi}{\Lambda}$, hence satisfy the Helmholz equation

$$(3') \qquad \Delta\Gamma = -\left(\frac{2\pi}{\Lambda}\right)^2 \Gamma$$

and we already pointed out that adding rotation invariance (and normalizing $\Gamma(0)$ to be 1) determines Γ to be the Dirac distribution on the circle of radius $\frac{2\pi}{\Lambda}$. Now the space $\mathcal{E}_\Lambda$

of all smooth maps φ satisfying $\Delta\varphi = -\left(\frac{2\pi}{\Lambda}\right)^2 \varphi$ has the following properties:
– if φ is in $\mathcal{E}_\Lambda$, then $g \cdot \varphi : x \mapsto \varphi(g^{-1}x)$ is in $\mathcal{E}_\Lambda$ for any $g \in E(2)$; this means $\mathcal{E}_\Lambda$ is an invariant subspace of the set of smooth maps;
– $\mathcal{E}_\Lambda$ has itself no closed invariant subspace if one uses the usual smooth topology for it: indeed if φ is any nonzero element in $\mathcal{E}_\Lambda$, it may be shown that the family of maps $g \cdot \varphi$, $g \in G$, generates a dense subspace of $\mathcal{E}_\Lambda$. perhaps a word of caution is useful here: while Γ is rotation-invariant and determines a G-invariant random field, it is certainly not itself invariant under the full group G of motions.

Let us insist that condition (3') may now be rewritten as:

(3") Γ belongs to one of the elementary invariant subspaces $\mathcal{E}_\Lambda$.

Let's then start with any square-integrable map f from $\mathbb{R}^2$ to $\mathbb{C}$ with continous Fourier transform; for each $K > 0$, we may restrict $\widehat{f}$ to the circle of radius K to produce the map

$$f_K = \vec{x} \mapsto \int_{\mathbb{S}^1} \widehat{f}(K\vec{u})e^{iK\vec{u}\cdot\vec{x}}d\vec{u}$$

which is automatically smooth, but not square-integrable [6]; see [64] for details. And then for almost every x,

$$f(x) = \int_{\mathbb{R}^+} f_K(x)KdK.$$

This shows that the $\mathcal{E}_\Lambda$ do provide the factors in the Plancherel decomposition of $\mathbb{L}^2(\mathbb{R}^2)$ described at the end of section 2.4.2, and the equivalence between conditions (3) and (3") shows how the spectral thinness condition found in models is related to the Plancherel decomposition of $\mathbb{L}^2(\mathbb{R}^2)$. In section 2.2, we saw how each of this factors determines a unique Gaussian Random Field which provides realistic V1-liks maps.

We now have gathered all the ingredients for building two-dimensional V1-like maps with non-Euclidean symmetries. But before we leave the Euclidean setting, let us remark that the irreducible representation of $SE(2)$ carried by the $\mathcal{E}_\Lambda$ has been used in [22], although the presentation there is rather different [7]. While the approach of [22], which brings the horizontal connectivity to the fore and uses Heisenberg's uncertainty principle to exploit the noncommutativity of $SE(2)$, has notable differences with using Gaussian random fields, it is very interesting and defines *real*-valued random fields which are good candidates for the maps a_ϑ of section 2.1. To the author's knowledge this is the first time irreducible representations of $SE(2)$ were explicitly used to study V1, and reading this paper was the starting point for the present study.

6. To be precise, we can multiply $\widehat{f}$ with the Dirac distribution on the circle of radius K, obtaining a tempered distribution on $\mathbb{R}^2$, and define f_K as the inverse Fourier transform of this multiplication: it is automatically a smooth function.

7. Since Fourier transforms of all maps in $\mathcal{E}_\Lambda$ are supported on a circle, we may see a function in any of the $\mathcal{E}_\Lambda$ as a complex-valued function on the unit circle; but the G-action depends on Λ. In this picture, any element $g = (R, \vec{x})$ of the Euclidean group gives an operator $T_\Lambda(g)$ on $\mathbb{L}^2(\mathbb{S}^1)$:

$$T_\Lambda(g)\Phi = \vec{u} \mapsto e^{i\frac{2\pi}{\Lambda}\vec{u}\cdot\vec{x}}\Phi(R^{-1}\vec{u});$$

and the Poisson transform $\Phi \in \mathbb{L}^2(\mathbb{S}^1) \mapsto \int_{\mathbb{S}^1} \widehat{f}(K\vec{u})e^{K\vec{u}\cdot\vec{x}}d\vec{u} \in \mathcal{E}_\Lambda$ is a continuous bijection that intertwines the representation T_Λ of $SE(2)$ with the natural representation on $\mathcal{E}_\Lambda$.

3 Results

3.1 Hyperbolic geometry

Let us now turn to plane hyperbolic geometry. The relevant groups for capturing the global properties of the hyperbolic plane assemble in the Klein pair $(G, K) = (SU(1,1), SO(2))$ as described in the Methods section.

If we are to look for pinwheel-like arrangments lurking in the representation theory of $SU(1,1)$, we need a familiarity with some irreducible representations. We shall use the next paragraph to give the necessary details on the geometry of the unit disk; the description of all unitary representations of $SU(1,1)$, however, we shall skip over [8] in order to focus on the Plancherel decomposition of $\mathbb{L}^2(G/K)$.

We should note at this point that hyperbolic geometry and SL_2-invariance [9] have been used by Chossat and Faugeras for a different purpose [39]; the same basic ingredients will appear here.

3.1.a Geometrical preliminaries

In this subsection, we must ask again for a little mathematical patience from our reader while we introduce some geometrical notions which we shall need for building hyperbolic maps (this is very standard material again; see [45], section 0.4, and the paper by Chossat and Faugeras). So let us first describe some further interplay between the algebraic structure of $G = SU(1,1)$ and hyperbolic geometry in the unit disk. Geodesics in $\mathbb{D}$ are easily described in terms of groups: since the action of G leaves the metric η invariant, the energy functional whose extremal paths are the geodesics of $\mathbb{D}$ is G-invariant as well; so any element $g \in G$ sends geodesics to geodesics. What is more, the horizontal path $t \mapsto \gamma(t) = (tanh(t), 0)$ has hyperbolic unit speed and it is not difficult to show that it is a geodesic of $\mathbb{D}$ (see [45], p. 29). Now, the interplay between group theory and Riemannian geometry makes it easy to find all geodesics of $\mathbb{D}$. Since $\gamma(t)$ is where the origin is sent by the element $\begin{pmatrix} \cosh(t) & \sinh(t) \\ \sinh(t) & \cosh(t) \end{pmatrix}$ of G, there is a subgroup of G to tell the story of the point 0 along this path: it is the subgroup

$$A = \left\{ \begin{pmatrix} \cosh(t) & \sinh(t) \\ \sinh(t) & \cosh(t) \end{pmatrix} \mid t \in \mathbb{R} \right\}.$$

From conjugates of this subgroup we may describe all geodesics in $\mathbb{D}$: if we start with a point x_0 of $\mathbb{D}$ and a tangent vector v_0 at x_0, there is an element g of G which sends both x_0 to the origin and v_0 to the right-pointing horizontal unit vector. But now the geodesic emanating from x_0 with speed v_0 is none other than the orbit of x_0 under the subgroup $g^{-1}Ag$; as g acts through a homography on $\mathbb{D}$, it is easy to see that this orbit draws on $\mathbb{D}$ a circle that is tangent to $x_0 + \mathbb{R}v_0$ and orthogonal to the boundary of $\mathbb{D}$ (this "circle" might be a line, which we can think of as a circle of infinite radius here).

Just as a family of parallel lines in $\mathbb{R}^2$ has an associated family of parallel hyperplanes that are orthogonal to each line in the family, the set of A-orbits has an associated family of *parallel horocycles*: writing b_0 for the point of the boundary $\partial\mathbb{D}$ that is in the closure of

8. For completeness we recall that there are unitary irreducible representations of $SU(1,1)$ which do not enter the Plancherel decomposition of $\mathbb{L}^2(G/K)$; the deep and beautiful work by Bargmann, Harish-Chandra on these representations will not appear in this paper.

9. The groups $SU(1,1)$ and $SL_2(\mathbb{R})$ are famously isomorphic, see for instance [39].

every A-orbit (i.e. the point 1+0i in $B = \partial\mathbb{D}$), a circle that is tangent to $\partial\mathbb{D}$ at b_0 meets every A-orbit orthogonally. What is more, given two such circles, there is on any A-orbit a unique segment that meets them both orthogonally; the length of this hyperbolic geodesic segment does not depend on the A-orbit chosen, so it is very reasonable indeed to call our two circles parallel. Circles tangent to $\partial\mathbb{D}$ were named *horocycles* by Poincaré, so we have been looking at the (parallel) family of those horocycles that are tangent to $\partial\mathbb{D}$ at b_0.

Now these horocycles too come with a group to tell their tale: they are the orbits in $\mathbb{D}$ of

$$N = \left\{ \begin{pmatrix} 1+is & -is \\ is & 1-is \end{pmatrix} \mid s \in \mathbb{R} \right\}.$$

To describe the families of parallel horocycles associated to other families of geodesics it is a conjugate of N that should be used: for this we should first note that if g is any element of the group, the family of $g^{-1}Ng$-orbits consists of horocycles tangent to $\partial\mathbb{D}$ at the same point, and then that each of these horocycles meets every $g^{-1}Ag$-orbit orthogonally.

There is one more definition that we shall need: it is closely linked to an important theorem in the structure of semisimple Lie groups [44, 53].

Theorem: Every element $g \in G = SU(1,1)$ may be written uniquely as a product [10] kan, where $k \in K$, $a \in A$ and $n \in N$. This is known as an *Iwasawa decomposition* for G.

Note that K, A and N are one-dimensional subgroups of G, but that the existence of a unique factorization $G = KAN$ does not mean at all that G is isomorphic with the direct product of K, A and N.

Even if this is a very famous result, an idea of the proof will be useful for us. Note first that any point $x \in \mathbb{D}$ may be reached from O by following the horizontal geodesic for a while (forwards or backwards) until one reaches the point of the horizontal axis which is on the same horocycle in the family of N-orbits, then going for x along this circle; this means that we can write $x = n \cdot (a \cdot O)$, where $n \in N$ and $a \in A$; now if g is any element of G, we may consider $x = g \cdot O$ and write it $x = n \cdot (a \cdot O) = (na) \cdot O$; then $(na)^{-1}g$ sends O to O, so it is an element of K. This proves the existence statement; uniqueness is easy but more technical.

Now if b is a point of the boundary $\mathbb{D}$ that has principal argument θ, we may view it as an element of K by assigning to it the element $\begin{pmatrix} e^{i\theta} & 0 \\ 0 & e^{-i\theta} \end{pmatrix}$; note that this element acts on $\mathbb{D}$ as a rotation of angle 2θ ! So beware, diametrally opposite elements b and $-b$ of the boundary define the same rotation.

If x is any point of $\mathbb{D}$ and b is any boundary point, we can now define a real number $\langle x, b\rangle$ as follows: start with any element $\tilde{x}$ of G that sends x to O, then choose a representative $\tilde{b}$ of b in K and consider the Iwasawa decomposition of the element $\tilde{x} \cdot \tilde{b}$ of G: it reads

$$\tilde{x} \cdot \tilde{b} = kan.$$

Now look at a, and consider the real number t such that $a = \begin{pmatrix} \cosh(t) & \sinh(t) \\ \sinh(t) & \cosh(t) \end{pmatrix}$. Write $\langle x, b\rangle$ for this number t; it is not difficult to check that this does not depend on any of the choices one has to make to select $\tilde{x}$ or $\tilde{b}$.

10. This is a product of matrices !

The indications we gave for the proof of the Iwasawa decomposition led Helgason to call $\langle x, b\rangle$ a (signed) *composite distance*; the definition and its relationship with the hyperbolic distance are illustrated on figure 3.

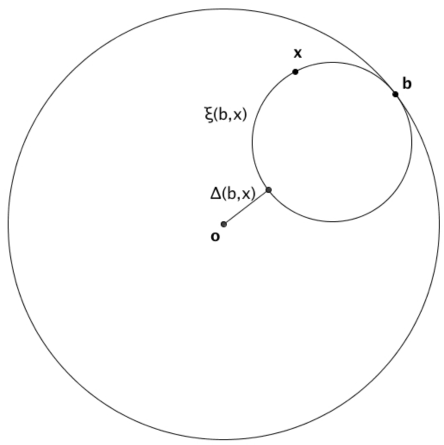


Figure 3: **The "composite distance" to a point of the boundary.** Definition of the quantity $\langle x, b\rangle$ if x is a point of $\mathbb{D}$ and b a point of its boundary: $\xi(b, x)$ is the horocycle through x which is tangent to the boundary at b, and $\Delta(b, x)$ is the segment joining the origin O to the point on $\xi(b, x)$ which is diametrically opposite b; the number $\langle x, b\rangle$ is, up to a sign, the hyperbolic length of this segment.

3.1.b Helgason waves and harmonic analysis

At first sight, there is no reason why harmonic analysis on the hyperbolic plane should "look like" Euclidean harmonic analysis: their invariance groups are apparently quite different and there is nothing like an abelian "hyperbolic translation group" whose characters may obviously be taken as a basis for building representation theory. So it may come as a surprise that there *are* analogues of plane waves in hyperbolic space, and (more importantly) that these enjoy much the same relationship to hyperbolic harmonic analysis as Fourier components do to Euclidean analysis. The discovery of these plane waves can be traced back to the seminal work of Harish-Chandra [51] on spherical functions of semi simple Lie groups (we shall come back to this in a moment), and their systematic use in non-euclidean harmonic analysis is due to Helgason [45, 46]. Since they will be a key ingredient in the rest of this section, let us now describe these waves.

Start with a point b of the boundary $B := \partial\mathbb{D}$ and a real number ω. For $z \in \mathbb{D}$, set

$$e_{\omega,b}(z) := e^{(i\omega+1)\langle z,b\rangle}.$$

This is a complex-valued function on $\mathbb{D}$; note that as z draws close to b, $\langle z, b\rangle$ goes to infinity, so $e_{\omega,b}(z)$ grows exponentially; on the other hand it decreases exponentially as z draws close to $-b$. This growth factor in the modulus is there for technical reasons, but has important consequences for representation theory and in the case of our orientation maps, it will have a clear influence on the pinwheel density we shall calculate later; we shall discuss this at the end of the present section. For a plot indicating the argument of $e_{\omega,b}$, see figure 4.

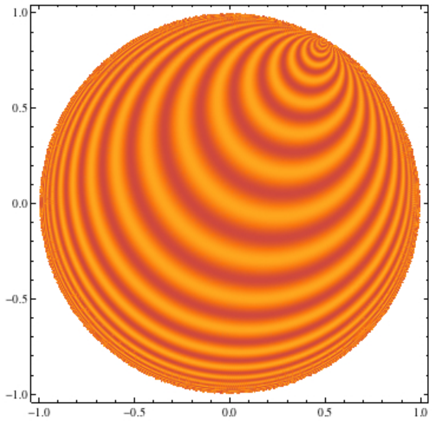


Figure 4: Plot of the real part of a Helgason wave, with the exponential growth factor deleted: in the darkest regions the real part of the scaled wave vanishes, and in the brightest regions it is equal to one. Given the formula for $e_{\omega,b}$, this plot also gives an idea of the argument as a function of z; notice that the argument is periodic when restricted to any geodesic whose closure in $\mathbb{C}$ contains the point of -1 of B.

Just as plane waves are generalized eigenvectors for the euclidean laplacian on $\mathbb{R}^2$, Helgason waves are "eigenvectors" for the relevant laplacian. Define, for $\mathcal{C}^2$ f,

$$\Delta_{\mathbb{D}} f := p \mapsto \frac{1}{\eta(p)} (\Delta_{\mathbb{R}^2} f)(p).$$

This is indeed the Laplace operator for $\mathbb{D}$: it can be defined from group-theoretical analysis alone, in much the same way we obtained the Poincaré metric in section 2.3 and the appendix. Theoretical questions aside, the reader may check easily that this new laplacian is G-invariant, that is, $\Delta_{\mathbb{D}} \left[f(g^{-1} \cdot) \right] = \left[\Delta_{\mathbb{D}} f \right] (g^{-1} \cdot)$.

Now, a crucial observation is that $e_{\lambda,b}$ is an eigenfunction for this operator, with a *real* eigenvalue [11] :

$$\Delta_{\mathbb{D}} e_{\lambda,b} = -(\omega^2 + 1) e_{\lambda,b}.$$

As a consequence, any finite combination of the $e_{\omega,b}$ with ω fixed is an "eigenvector" for $\Delta_{\mathbb{D}}$. Now let us go for continous combinations: if $\mu : B \to \mathbb{R}$ is a continuous function, then

$$\mathcal{P}_{\omega}(\mu) := z \mapsto \int_B e_{\omega,b}(z) \mu(b) db$$

(the *Poisson transform* of μ) is another eigenfunction with eigenvalue $-(\omega^2 + 1)$:

$$\Delta_{\mathbb{D}}[\mathcal{P}_{\omega}(\mu)] = -(\omega^2 + 1)[\mathcal{P}_{\omega}(\mu)].$$

We shall write $\mathcal{E}_{\omega}(\mathbb{D})$ for the space of all such smooth eigenfunctions:

$$\mathcal{E}_{\omega}(\mathbb{D}) := \left\{ f \in \mathcal{C}^{\infty}(\mathbb{D}, \mathbb{C}) \mid \Delta_{\mathbb{D}} f = -(\omega^2 + 1) f \right\}.$$

For each continuous function $\mu : B \to \mathbb{R}$, we then know that $\mathcal{P}_{\omega}(\mu)$ belongs to $\mathcal{E}_{\omega}(\mathbb{D})$, and in fact the image of $\mathcal{P}_{\omega}$ is dense in $\mathcal{E}_{\omega}(\mathbb{D})$ for several natural topologies (see [45], chapter 0, Theorem 4.3, Lemma 4.20). Since $\Delta_{\mathbb{D}}$ is G-invariant, $\mathcal{E}_{\omega}(\mathbb{D})$ is a stable subspace

11. That the eigenvalue should be real is the technical reason why the growth factor in the modulus is needed.

of $\mathcal{C}^\infty(\mathbb{D})$; by studing $\mathcal{P}_\omega$, Helgason was able to prove that the $\mathcal{E}_\omega(\mathbb{D})$ is irreducible ([45], chapter 0, Theorem 4.4). The following theorem then achieves the Plancherel decomposition of $\mathbb{L}^2(\mathbb{D})$ in the sense of section 2.4.2, and is a cornerstone of harmonic analysis on the unit disk (see [45], chapter 0, Theorem 4.2; the extension to $\mathbb{L}^2$ is proved there also) :

Theorem (Harish-Chandra, Helgason): for each $f \in \mathcal{C}^\infty(\mathbb{D})$, write $f_\omega(z) = \int_B \left(\int_{\mathbb{D}} f(y)e_{-\omega,b}(y)dy\right) e_{\omega,b}(z)db$, and set $\Pi(\omega) = \frac{\omega}{2}\tanh(\frac{\pi\omega}{2})$ for each positive ω; then the following equality holds as soon as all terms are defined by converging integrals:

$$f(z) = \int_{\mathbb{R}^+} f_\omega(z)\ \Pi(\omega)d\omega.$$

When we build our hyperbolic maps in the next section, the $\mathcal{E}_\omega(\mathbb{D})$ are the only representations we shall need. We will come back to this shortly.

3.1.c Hyperbolic orientation maps

It is time to build our hyperbolic analogue of Orientation Preference Maps. Suppose we wish to arrange sensors on $\mathbb{D}$ so that each point of $\mathbb{D}$ is equipped with a receptive profile which has an orientation preference and a selectivity. This may be local model for an arrangement of $V1$-like receptive profiles on a negatively curved region of the cortical surface, and though its primary interest is probably in clarifying the role of symmetries in discussions, the construction to come can be thought of in this way.

We shall require that this arrangment have the same randomness structure (condition (1)) as the Euclidean model of the Methods section, that is, be a "typical" realization of a standard complex-valued gaussian Random Field on the space $\mathbb{D}$, say $\mathbf{z}$. If it is to have an analogous invariance structure (conditions (2) and (3)), it should, first, be assumed to be G-invariant; what is more, we should look for a field that probes an irreducible factor of the representation of G on $\mathbb{L}^2(\mathbb{D})$ (see section 2.4.2); as a consequence, any realization of $\mathbf{z}$ should be an eigenfunction of $\Delta_{\mathbb{D}}$, with the eigenvalue determined by $\mathbf{z}$. Remembering the Euclidean terminology we used in the Methods section, let us make a

Definition: *a monochromatic gaussian field on* $\mathbb{D}$ is a complex-valued gaussian random field on $\mathbb{D}$ whose probability distribution is $SU(1,1)$-invariant and which takes values in one of the $\mathcal{E}_\omega$, $\omega > 0$. If $\mathbf{z}$ is such a field, the positive number ω will be called the spectral parameter of $\mathbf{z}$.

To see how to build such a monochromatic field, we should translate our requirements into a statement about its covariance function; luckily there is a theorem here ([34], see the discussion surrounding Theorem 6' and Theorem 7, in particular eq. (3.20) there) that says our conditions on $\mathbf{z}$ are fulfilled if, and only if, the covariance function of $\mathbf{z}$, when turned thanks to the G- invariance of $\mathbf{z}$ into a function from $\mathbb{D}$ to $\mathbb{C}$, is an elementary spherical function for $\mathbb{D}$ (a radial function on $\mathbb{D}$ which is an eigenfunction $\Delta_{\mathbb{D}}$). What does this mean ?

First, note that the covariance function $C : \mathbb{D}^2 \to \mathbb{C}$ of our field may be seen as a function $\tilde{C} = G^2 \to \mathbb{C}$: we need only set $\tilde{C}(g_1, g_2) = C(g_1 \cdot O, g_2 \cdot O)$. Now, that $\mathbf{z}$ should be G-invariant means that for every $g_0 \in G$, $\tilde{C}(gg_1, gg_2)$ should be equal to $C(g_1, g_2)$; in particular, writing $\Gamma(g)$ for $\tilde{C}(g, 1_G)$, we get $\tilde{C}(g_1, g_2) = \Gamma(g_2^{-1} \cdot g_1)$. The whole of the correlation structure of the field is summed up in this Γ, which is a function from G to $\mathbb{C}$.

Now, not every function from G to $\mathbb{C}$ can be obtained in this way: since it should come from a function C which is defined on $\mathbb{D}^2$ and thus satisfies $C(g_1k_1, g_2k_2) = C(g_1, g_2)$ when k_1, k_2 is in K, it should certainly satisfy $\Gamma(k_1gk_2) = \Gamma(g)$ for $k_1, k_2 \in K$; so Γ does in fact define a function on $\mathbb{D}$ and this function is left-K-invariant, that is, radial in the usual sense of the word (*property (A)*). What is more, since the field is assumed to have variance 1 everywhere, it should also satisfy *(B)* $\Gamma(\mathrm{Id}_G) = 1$ (*property (B)*).

Let us now add that monochromaticity for $\mathbf{z}$ is equivalent to Γ being an eigenfunction of $\Delta_{\mathbb{D}}$ (*property (C)*).

Functions on $\mathbb{D}$ with properties *(A), (B)* and *(C)* are called elementary spherical functions for $\mathbb{D}$.

Now, we stumbled upon these (following Yaglom) while looking for pinwheel-like structures, but spherical functions (and their generalizations to semisimple symmetric spaces) have been intensely studied in the last half of a century. In fact, they were defined by Elie Cartan as early as 1929 with the explicit objective of determining the irreducible components of $\mathbb{L}^2(G/K)$ for a large class of Klein pairs (G, K). The following theorem will look like an easy consequence of everything we discussed earlier, but history went the other way and it is in looking for spherical functions that Harish-Chandra discovered what we called Helgason waves.

Theorem (Harish-Chandra 1958, [51]): In each of the irreducible components $\mathcal{E}_\omega(\mathbb{D})$, there is a unique spherical function; it is the map

$$\varphi_\omega := x \mapsto \int_B e_{\omega,b}(x)db.$$

If we plot φ_ω it will resemble the Euclidean Bessel-kind covariance functions; only there is a marked difficulty in dealing with the growth at infinity of these functions, which accounts for some (not all, of course) of the many difficulties Harish-Chandra and Helgason had to overcome in developing harmonic analysis on $\mathbb{D}$.

The properties of elementary spherical functions include the conditions which guarantee, thanks to the existence theorem by Kolmogorov mentioned in the Methods section, that each of the φ_ω really is the covariance function of a Gaussian field on $\mathbb{D}$. So we can summarize the preceding discussion with the following statement:

Proposition A: For each $\omega > 0$, there is exactly one monochromatic gaussian field on $\mathbb{D}$ with spectral parameter ω.

These are our candidates for providing V1-like maps on $\mathbb{D}$. We now need to see, by plotting one, whether a “typical" sample of a monochromatic field looks like a hyperbolic V1-like map, hence we need to go from the covariance function to a plot of the field itself. All technical details aside, Euclidean and hyperbolic spherical functions are close enough for the transition from a spherical function to the associated gaussian field to be exactly the same in both cases. We said nothing of this step in the Methods section, so let us come back to the Euclidean setting for a second.

Spherical functions there we already described: they are Fourier transforms of the Dirac distribution on a circle, so they read

$$\psi_R := x \mapsto \int_B e^{iRb\cdot x}db.$$

Now, this builds ψ_R out of a constructive interference between plane waves $e_{R,b}$, $b \in B$. In order to obtain a gaussian random field while keeping the eigenfunction property, we need only attribute a random gaussian weight (which is a complex number, this includes phase) to each of our plane waves. This needs some care since we are dealing with a continuum of weights to attribute, but there is a standard random measure $\mathbb{Z}$ on the circle, the standard Gaussian white noise, which is meant to achieve this (see the appendix for details, and also [66]): this produces an invariant random field

$$\mathbf{z}_R := x \mapsto \int_B e^{iRb\cdot x} d\mathbb{Z}(b)$$

whose covariance is ψ_R as desired. We give further details on the transition from ψ_R to $\mathbf{z}_R$ in the appendix.

This construction depends only on properties which are common to the Euclidean and hyperbolic plane; thus, it transfers unimpaired to the hyperbolic plane.

We know at last what an orientation preference-like map should look like in hyperbolic geometry: pick a positive number ω; out of the stochastic integral

$$x \mapsto \int_{\mathbb{S}^1} e_{\omega,b}(x) d\mathbb{Z}(b)$$

there will arise orientation maps. Just as in the euclidean case, they are readily approximated by picking a number of regulary spaced points $b_1, \ldots b_n$ on the boundary circle, assigning them independent reduced gaussian weights $\zeta_1, \ldots \zeta_n$ in $\mathbb{C}$ (so the ζ_i are complex-valued reduced gaussian random variables, independent from each other) and considering

$$x \mapsto \frac{1}{n} \sum_{k=1}^{n} \zeta_i e_{\omega,b_i}(x).$$

A computer-generated sample is shown on figure 5; comparison with Escher's celebrated drawings of periodic tilings of $\mathbb{D}$ [59] might be telling.

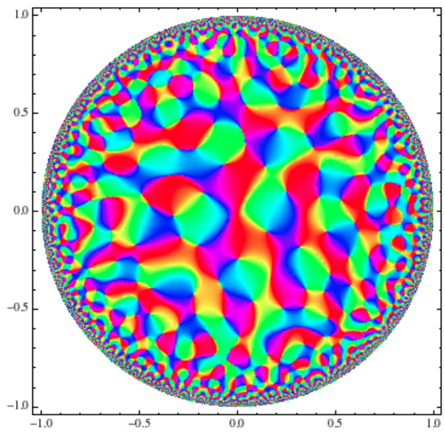


Figure 5: Plot of a monochromatic "orientation map" on the hyperbolic plane. We used the spectral parameter $\omega = 18$ in units of the disk's radius. Because of the growth factor in the modulus of the $e_{\omega,b}$, drawing a picture in which discretization effects do not appear calls for using more propagation directions than it did in the Euclidean case: 200 directions were used to generate the drawn picture.

So there does appear a quasi-periodic tiling of the unit disk; it should not of course be forgotten that this quasi-periodicity holds only when the area of an "elementary cell" is measured in the appropriate hyperbolic units (see the previous section and the next).

3.1.d Hyperbolic pinwheel density

What is this area σ of an elementary cell, by the way, and can we estimate the density of pinwheels per area σ ?

In the Euclidean case, we used results from physics that originally dealt with superpositions of Euclidean waves. Of course singularities in superpositions of random waves do occur in many interesting physical problems: interest first came from the study of waves travelling through the (irregular) arctic surface; quantum physics has naturally been providing many interesting random superpositions: they occur in laser optics [61], superfluids [63]... This has prompted recent mathematical developments. In this section, we would like to point out that these are now sharp enough to allow for calculations *outside* Euclidean geometry.

Consider an invariant monochromatic random field $\mathbf{z}$, and write ω for the corresponding "wavenumber" (so that $\mathbf{z}$ belongs to $\mathcal{E}_\omega$). We would like to evaluate the expectation for the number of pinwheels (zeroes of $\mathbf{z}$) in a given domain $\mathcal{A}$ of the unit disk. Let us write $\mathcal{N}_\mathcal{A}$ for the random variable recording the number of pinwheels in $\mathcal{A}$. We will now evaluate the expectation of this random variable, and the result will be summarized as Theorem A below.

Since $\mathbf{z}$ is G-invariant, it is to be expected that $\mathbb{E}\left\{\mathcal{N}_\mathcal{A}\right\}$ depends only on the *hyperbolic* area of $\mathcal{A}$: our first claim is that this is indeed the case: writing $|\mathcal{A}|_h$ for the hyperbolic area of $\mathcal{A}$, let us show that

$$\frac{\mathbb{E}\left\{\mathcal{N}_\mathcal{A}\right\}}{|\mathcal{A}|_h} = \frac{V_0}{\pi}$$

where V_0 stands for the variance of the real-valued random variable $\partial_x \mathfrak{Re}(\mathbf{z})(0)$. For this we shall use Azaïs and Wschebor's version of the Kac-Rice formula for random fields, which in our setting says:

Theorem (see [25], Theorem 6.2): assume $\mathbf{z}$ is a smooth, reduced [12] Gaussian random field from $\mathbb{D}$, which almost surely has no degenerate zero [13] in $\mathcal{A}$; then

$$\mathbb{E}\left\{\mathcal{N}_\mathcal{A}\right\} = \frac{1}{2\pi}\int_\mathcal{A} \mathbb{E}\left\{|\det\, d\mathbf{z}(p)|\,|\; \mathbf{z}(p)=0\right\} dp$$

(here the integral is Lebesgue integral, and the integrand is a conditional expectation).

To use this theorem, we should note (see [Adl]) that in a field with constant variance, at each point p the value any derivative of any component of the field is independent (as a random variable) from the value of the field at p; so the two variables $\mathrm{Det}[d\mathbf{z}](p)$ and $\mathbf{z}(p)$ are independent too; thus for invariant fields on $\mathbb{D}$ we know that the hypotheses in the theorem are satisfied, and that in addition we may remove the conditioning in the expectation formula. So we are left with evaluating the mean determinant of a matrix whose columns are independent gaussian vectors, with zero mean and the same variance V_p

12. This means each $\mathbf{z}(x)$ is a complex-valued Gaussian random variable with zero mean and variance one.

13. A degenerate zero is a point at which both $\mathbf{z}$ and $d\mathbf{z}$ are zero.

as $\partial_x \mathfrak{Re}(\mathbf{z})(p)$. We are left with evaluating the Euclidean area of the random parallelogram generated by these random vectors, and using the "base times height" formula it is easy to prove this mean area is $2V_p$.

So we need to see that $\displaystyle\int_{\mathcal{A}} \frac{V_p}{\pi} dp$ is equal to $|\mathcal{A}|_h \dfrac{V_0}{\pi}$. But this is easy: when the real-valued Gaussian field $\zeta = \mathfrak{Re}\mathbf{z}$ is G-invariant, we can define a G-invariant riemannian metric on $\mathbb{D}$ by setting $g^{\zeta}_{ij}(p) = \mathbb{E}\left\{\partial_i \zeta(p) \partial_j \zeta(p)\right\}$; as we said in the Methods section, this must be a constant multiple of the Poincaré metric. It follows that V_p is equal to $\eta(p)V_0$, while $\eta(p)$ is the hyperbolic surface element. This proves the announced formula $\mathbb{E}\left\{\mathcal{N}_{\mathcal{A}}\right\} = |\mathcal{A}|_h \frac{V_0}{\pi}$.

Now, evaluating the variance of the first derivative $\partial_x \mathfrak{Re}(\mathbf{z})(0)$ is easy: it is obtained from the *second* derivative with respect to the x-coordinate [14] of the covariance function Γ of the random field $\mathfrak{Re}(\mathbf{z})$,

$$\mathbb{E}\left\{(\partial_1 \mathfrak{Re}(\mathbf{z})(0))^2\right\} = \partial_{1,x_1}\partial_{1,x_2}\mathbb{E}\left\{\mathfrak{Re}(\mathbf{z})(x_1)\mathfrak{Re}(\mathbf{z})(x_2)\right\}|_{x_1=x_2=x} = \partial_{1,x}\partial_{1,y}\Gamma(\alpha(x)\alpha(y)^{-1})\big|_{x=y=0},$$

where α is a smooth section of the projection from G to $\mathbb{D}$ induced by the action of the origin (such a smooth section does exist). For a G-invariant field $\partial_{1,x}\partial_{1,y}\Gamma(\alpha(x)\alpha(y)^{-1})\big|_{x=y=0} = \partial_{2,x}\partial_{2,y}\Gamma(\alpha(x)\alpha(y)^{-1})\big|_{x=y=0}$; as a consequence V_0 is half the value of $-\Delta\Gamma$ at zero. Now $\Delta\Gamma = -(\omega^2+1)\Gamma$ and $\Gamma(0) = 1/2$ (the value of the covariance function of all of $\mathbf{z}$ at zero is one, but here we are dealing only with the real part), so we have obtained the following result:

Theorem A: *suppose* $\mathbf{z}$ *is the only complex-valued, centered Gaussian Random field on* $\mathbb{D}$ *whose probability distribution is* $SU(1,1)$*-invariant, and whose correlation function, when turned into a function on* $SU(1,1)$*, is Harish-Chandra's spherical function* φ_ω*. Consider a Borel subset* $\mathcal{A}$ *of* $\mathbb{D}$*, write* $|\mathcal{A}|_h$ *for its area w.r.t the Poincaré metric, and* $\mathcal{N}_{\mathcal{A}}$ *for the random variable recording the number of zeroes of* $\mathbf{z}$ *in* $\mathcal{A}$*. Then*

$$\frac{\mathbb{E}\left\{\mathcal{N}_{\mathcal{A}}\right\}}{|\mathcal{A}|_h} = \frac{\pi}{\omega^2+1}.$$

It is worth pointing out that the proof above no longer features any reference to wave propagation; we just needed the invariance properties of our covariance function and a nice property of our new laplacian. This means that calculations should travel unimpaired to geometries where nothing like wave propagation is available for building spherical functions and representation theory. We shall see this at work on the sphere in the next section.

But let us linger a moment in the hyperbolic plane, for our new monochromatic maps do exhibit a rather unexpected feature: while same-phase wavefronts in $e_{\omega,b}$ line up at hyperbolic distance $\frac{2\pi}{|\omega|}$, it seems like the right hyperbolic area for a “hyperbolic hypercolumn", that area which we called σ at the beginning of this subsection, should be $\frac{4\pi^2}{\omega^2+1}$. In fact, we claim that the typical hyperbolic distance between two points in the map that have the same orientation preference is *not* $\frac{2\pi}{|\omega|}$ as we would guess by thinking in Euclidean terms, but $\frac{2\pi}{\sqrt{\omega^2+1}}$. There is something of course to support of this claim: we can evaluate the typical spacing by selecting a portion of a geodesic and evaluate the mean number of

14. For legibility we rewrote the derivative in the horizontal direction as ∂_1 in the next formula: so if φ is a function of two variables $x_1, x_2 \in \mathbb{D}$, ∂_{1,x_1} denotes the derivative in the horizontal direction of $x_1 \mapsto \varphi(x_1, x_2)$.

points with a given orientation preference. To motivate the statement of our result, see the discussion preceding Theorem C below, and also [73].

Theorem B: *suppose* $\mathbf{z}$ *is the only complex-valued, centered Gaussian Random field on* $\mathbb{D}$ *whose probability distribution is* $SU(1,1)$*-invariant, and whose correlation function, when turned into a function on* $SU(1,1)$*, is Harish-Chandra's spherical function* φ_ω*. Select a geodesic* δ *on* $\mathbb{D}$*, consider a segment* Σ *on* δ*, and write* $|\Sigma|_h$ *for its hyperbolic length. Write* Ψ *for the real-valued random field on* Σ *obtained by projecting the values of* $\mathbf{z}|_\Sigma$ *onto an arbitrary axis in* $\mathbb{C}$*, and* $\mathcal{N}_\Sigma$ *for the random variable recording the number of zeroes of* Ψ *on* Σ*. Define* Λ *as the only positive number such that* $\Delta_{\mathbb{D}}\mathbf{z} = -\left(\frac{2\pi}{\Lambda}\right)^2 \mathbf{z}$*. Then*

$$\frac{\mathbb{E}[\mathcal{N}_\Sigma]}{|\Sigma|_h} = \frac{1}{\Lambda}.$$

Note that the zeroes considered here are points on Σ where the preferred orientation is the vertical, and have nothing to do with pinwheel centers (which were the zeroes considered in Theorem A).

Let us give a summarized proof of this result here (see also the discussion leading to Theorem C, where the arguments are similar but the idea appears perhaps more clearly). Set $\mathbf{u} := \mathfrak{Re}(\mathbf{z}|_\delta)$; this is a real-valued random field on the geodesic δ. Since δ is an orbit on $\mathbb{D}$ of a one-parameter subgroup of $SU(1,1)$, we can view it as a real-valued, stationary random field on the real line and apply the classical one-dimensional Kac-Rice formula. Using the one-parameter subgroup to transfer the result back to δ, and using the shift-invariance of $\mathbf{z}$, we get the following formula :

$$\frac{\mathbb{E}[\mathcal{N}_\Sigma]}{|\Sigma|_h} = \frac{\sqrt{\lambda_2}}{\pi}$$

where $\lambda_2 = \mathbb{E}[\mathbf{u}'(0)^2]$ is the second spectral moment of the field $\mathbf{u}$. But we actually evaluated $\sqrt{\lambda_2}$ while proving Theorem A: it is equal to $\frac{\omega^2+1}{\pi}$. This completes the proof of Theorem B.

3.2 Spherical geometry

Let us now examine the positively-curved case, *viz.* the sphere $\mathbb{S}^2$. Recall from the Methods section that the geometry of the sphere is captured by the Klein pair $(SO(3), SO(2))$.

We will start by looking for an orientation preference-like map on the sphere. Let us therefore look for an arrangment $\mathbf{z}$ with our usual randomness structure, that is, for a complex-valued standard Gaussian random field on the space $\mathbb{S}^2$; let us further assume that the field $\mathbf{z}$ is G-invariant and probes an irreducible factor of the natural representation of $SO(3)$ on $\mathbb{L}^2(\mathbb{S}^2)$ (see 2.4.2). The arguments we used for the hyperbolic plane go through, so we are now looking for a gaussian Random Field whose covariance function is an elementary spherical function for $\mathbb{S}^2$.

In the last section, we built these out of hyperbolic analogues of Euclidean plane waves; here there is no obvious "plane wave" candidate for carrying the torch. However, it is quite easy to find alternative building-blocks for the irreducible factors of the representation of G on $\mathbb{L}^2(\mathbb{S}^2)$: these are the familiar *spherical harmonics*; since there will be a significant difference between the maps we shall describe and those we encountered on non-positively curved spaces of the preceding sections, we shall take the next paragraph to examine their rôle in representation theory even if this is famous textbook material, see [67], Chap. 7.

3.2.a Preliminaries on the spherical harmonics and the Plancherel decomposition of $\mathbb{L}^2(\mathbb{S}^2)$

The sphere has its own Laplace operator, just as the Euclidean plane and the hyperbolic plane do. To define it, regard $\mathbb{S}^2$ as isometrically embedded in $\mathbb{R}^3$ as the unit sphere centered at the origin O. If f is a smooth function on the sphere, we may extend it to a smooth function $\tilde{f}$ on $\mathbb{R}^3 - O$ that is constant on every ray issued from the origin: $\tilde{f}(x) = f(x/\|x\|)$. Define now

$$\Delta_{\mathbb{S}^2} f = \left(\Delta_{\mathbb{R}^3} \tilde{f}\right)|_{\mathbb{S}^2}.$$

Since $\Delta_{\mathbb{R}^3}$ is rotation-invariant, $\Delta_{\mathbb{S}^2}$ is rotation-invariant also; so every "eigenspace" of $\Delta_{\mathbb{S}^2}$ on $\mathbb{L}^2(\mathbb{S}^2)$ is a G-invariant subspace. To get eigenfunctions for $\Delta_{\mathbb{S}^2}$, we need only remark that if Y is a homogeneous function of degree $\ell+1$ on $\mathbb{R}^3$,

$$(\Delta_{\mathbb{R}^3} Y)\big|_{\mathbb{S}^2} = \ell(\ell+1) Y\big|_{\mathbb{S}^2} + \Delta_{s^2}(Y\big|_{\mathbb{S}^2})$$

(here ℓ is a nonnegative integer).

If we start with a homogeneous function $\Phi : \mathbb{R}^3 \to \mathbb{C}$ of degree $\ell+1$ that is in addition *harmonic*, which means that it satisfies $\Delta_{\mathbb{R}^3}\Phi = 0$, and restrict it to the sphere, we get an eigenfunction for $\Delta_{\mathbb{S}^2}$, with eigenvalue $\ell(\ell+1)$. Actually, any member of the corresponding eigenspace can be extended to a harmonic homogeneous function of degree $\ell+1$, so this describes the whole eigenspace. Now, it turns out that every member of this eigenspace extends to a harmonic homogeneous *polynomial* function of degree $\ell+1$! If we write $\mathcal{H}_\ell$ for the $\ell(\ell+1)$-eigenspace of $\Delta_{\mathbb{S}^2}$, this space is then finite-dimensional, and its dimension is readily seen to be $2\ell+1$. Being finite-dimensional, $\mathcal{H}_\ell$ is a closed subspace of $\mathbb{L}^2(\mathbb{S}^2)$, so the usual scalar product on $\mathbb{L}^2(\mathbb{S}^2)$ restricts to a scalar product on $\mathcal{H}_\ell$. Here do Laplace's spherical harmonics come into play, for they give an orthonormal basis for $\mathcal{H}_\ell$: if we use spherical coordinates (θ, ϕ) on $\mathbb{S}^2$ and define, for $\ell \in \mathbb{N}^\star$ and $m \in \{-\ell, \ell\}$,

$$Y_{\ell,m}(\theta,\phi) := e^{im\phi} P_\ell(\cos\theta)$$

where P_ℓ is the ℓ-th Legendre polynomial, then $\{Y_{\ell,-\ell}, ..., Y_{\ell,0}, ...Y_{\ell,\ell}\}$ is an orthonormal basis for $\mathcal{H}_\ell$.

We should add at this point that the natural representation of $SO(3)$ on $\mathcal{H}_\ell$ is indeed irreducible; in the next section we shall associate an orientation preference map to each of the $\mathcal{H}_\ell$. But before we close this section, let us see how this relates to the decomposition of $\mathbb{L}^2(\mathbb{S}^2)$.

Starting from an element $f \in \mathbb{L}^2(\mathbb{S}^2)$, we may produce a countable set of coefficients by setting, for each $\ell \in \mathbb{N}^\star$ and each $m \in \{-\ell, ..\ell\}$,

$$\widehat{f}(\ell,m) := \int_{\mathbb{S}^2} f(x) Y_{\ell,m}(x) dx.$$

For each value of ℓ, this yields an eigenfunction of Δ related to f, namely $f_\ell := x \mapsto \sum_m \widehat{f}(\ell,m) Y_{\ell,m}(x)$. We are thus defining a *projection operator* $\mathcal{P}_\ell : \mathbb{L}^2(\mathbb{S}^2) \to \mathcal{H}_\ell$ (note that $\mathcal{P}_\ell^2 = \mathcal{P}_\ell$).

Now, it is a very famous theorem of Hermann Weyl that the initial map f can be reconstructed from this generalized Fourier series:

$$f = \sum_{\ell \geq 0} f_\ell.$$

Here, convergence of the right-hand side is to be understood in the mean-quadratic sense; but if f is smooth, uniform convergence does hold.

Notice that if f were to be an eigenfunction of $\Delta_{\mathbb{S}^2}$ but were to belong to none of the $\mathcal{H}_\ell$, $\mathcal{P}_\ell f$ would be zero for every ℓ, and so would f: Weyl's theorem thus indicates that there is no other eigenvalue of $\Delta_{\mathbb{S}^2}$ (thus the Peter-Weyl theorem reduces to the spectral theorem for the hermitian operator $\Delta_{\mathbb{S}^2}$ in the special case considered here). Of course it also achieves the Plancherel decomposition of section 2.4.2,

$$\mathbb{L}^2(\mathbb{S}^2) = \bigoplus_{\ell \geq 0} \mathcal{H}_\ell.$$

Notice that all analytic difficulties in the decomposition have vanished (the irreducible factors $\mathcal{H}_\ell$ are really spaces of square-integrable functions), and that Fourier *series* are enough to reconstruct a function, which means here that a countable set of irreducible representations is enough to decompose $\mathbb{L}^2(\mathbb{S}^2)$. Recall from section 2.4.2. that there is a simple reason for the marked differences between what happens on the sphere and what happens in our previous examples: Hermann Weyl proved that when the group G is *compact*, there is but a countable set of equivalence classes of irreducible representations.

3.2.b Spherical orientation maps

We have now at our disposal everything that is needed for orientation preference-like maps on the sphere, and on top of it, one important observation: our set of spherical maps, unlike the set of its Euclidean or Hyperbolic analogues, is discrete in nature. Out of the spherical harmonics $Y_{\ell m}$ arises one irreducible factor of $\mathbb{L}^2(\mathbb{S}^2)$ per ℓ; we feel it is appropriate to name the corresponding invariant gaussian random field a *spin ℓ monochromatic field.*

In the Euclidean and hyperbolic cases, we got all the information from the covariance function of the field; here we can dispense with the covariance function and describe such a field, say Φ_ℓ, a bit more explicitly than we could do for the previous invariance structures. Since the representation space $\mathcal{H}_\ell$ is finite-dimensional, specifying an orthonormal basis $(Y_{\ell,m})_m$ for $\mathcal{H}_\ell$ easily yields a gaussian probability law on $\mathcal{H}_\ell$: we need only consider

$$\sum_{m=-\ell}^{\ell} \zeta_m Y_{\ell,m}$$

where the ζ_m are reduced complex-valued gaussian random variables independent from each other. Now $\mathcal{H}_\ell$ may be seen as a function space on $\mathbb{S}^2$, with the corresponding functions readily written

$$\Phi_\ell : x \mapsto \sum_{m=-\ell}^{\ell} \zeta_m Y_{\ell,m}(x).$$

Thus, a standard gaussian probability law on $\mathcal{H}_\ell$ defines a gaussian random field. Now G acts on $\mathcal{H}_\ell$ by unitary operators, and the probability density $\sum \zeta_m Y_{\ell,m}$ is $G-$ invariant; this means that

$$x \mapsto \sum_{m=-\ell}^{\ell} \zeta_m Y_{\ell,m}(x)$$

is an invariant gaussian field that spans the “spin ℓ" -irreducible subspace of $\mathbb{L}^2(\mathbb{S}^2)$. This is precisely our spin ℓ monochromatic field; it is the only G-invariant standard gaussian field with values in $\mathcal{H}_\ell$.

We have plotted a map sampled from this field on fig. 6.

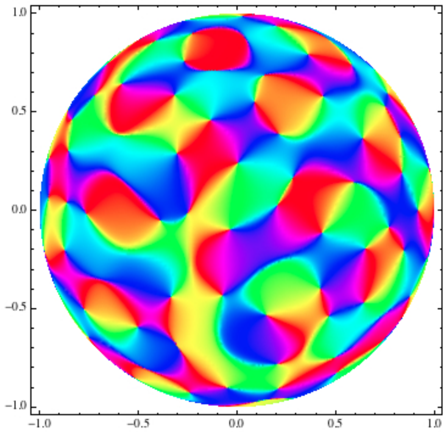


Figure 6: An orientation map on the sphere sampled from the argument of a monochromatic $SO(3)$-invariant gaussian random field on the sphere with spin 7. We plotted the restriction to a hemisphere; we used a superposition of spherical harmonics with spin seven and random, reduced independent gaussian weights.

3.2.c Spherical pinwheel density

Expectation values for pinwheel densities in spherical maps may be evaluated with the same methods we used in the previous sections. Here however, there appears a significant difference with the Euclidean and Hyperbolic cases: while monochromatic fields in those cases were indexed by a continuous parameter that is easily interpreted as a wavelength, there is apparently no natural scale for writing pinwheel densities.

In this subsection, we shall answer the following two questions:

(a) What is the mean (spherical) distance Λ between iso-orientation domains in a field that probes $\mathcal{H}_\ell$?

(b) What is the mean number of pinwheels within a given subset of the sphere, relative to the (spherical) area of this subset ? Is it $\frac{\pi}{\Lambda^2}$?

To answer the first question, let us select a geodesic segment on $\mathbb{S}^2$, that is, a portion of a great circle. What is, on this segment, the mean number of points where Φ_ℓ exhibits a given orientation ? Since standard gaussian fields are shift-invariant, we can consider a fixed value of the orientation, say the vertical. Points where Φ_ℓ exhibits this orientation are points where $\mathfrak{Re}(\Phi_\ell)$ vanishes; so let's define $\Psi_\ell = \mathfrak{Re}(\Phi_\ell)$ and look for its zeroes on the given great circle.

Now, Ψ_ℓ is a gaussian field on our great circle that is invariant under any rotation around this circle. This may be thought of as a *stationary* (translation-invariant) random field on $\mathbb{R}$ -an instance where the classical Kac-Rice formula ([30]) applies (think of what happens if one rolls this circle around on a Euclidean plane at constant speed). So we may assert that if $\mathcal{N}_\Sigma$ is the random variable recording the number of zeroes of Ψ_ℓ on Σ,

$$\frac{\mathbb{E}[\mathcal{N}_\Sigma]}{|\Sigma|} = \frac{\sqrt{\lambda_2}}{\pi}$$

where $\lambda_2 = \mathbb{E}[\Psi_\ell''(0)^2]$ is the second spectral moment of the field Ψ_ℓ. But now $\lambda_2 = \frac{\ell(\ell+1)}{4}$; if we set

$$\Lambda := \frac{2\pi}{\sqrt{\ell(\ell+1)}}$$

we have then obtained the following result.

Theorem C: *suppose Φ_ℓ is the only complex-valued, centered Gaussian Random field on $\mathbb{S}^2$ whose probability distribution is rotation-invariant, and whose samples belong to the irreducible subspace of $\mathbb{L}^2(\mathbb{S}^2)$ spanned by the spherical harmonics $Y_{\ell m}$, $m = -\ell, ..\ell$. Consider a geodesic segment Σ on $\mathbb{S}^2$, write $\mathcal{N}_\Sigma$ for the random variable recording the number of points on Σ where Φ_ℓ takes real values, and write Λ for the positive number $\frac{2\pi}{\sqrt{\ell(\ell+1)}}$. Then*

$$\frac{\mathbb{E}[\mathcal{N}_\Sigma]}{|\Sigma|} = \frac{1}{\Lambda}.$$

Thus, if the mean number of points on Σ to which Φ_ℓ attributes a given orientation preference is to be no less or no more than one, the length of Σ must be Λ. This answers question (a).

Now, if $\mathcal{A}$ is a subset of the sphere, denote by $|A|_s$ its spherical area and by $\mathcal{N}_\mathcal{A}$ the random variable recording the number of pinwheels of Φ_ℓ in A. Then, as in the previous cases, we observe a scaled density of π:

Theorem D: *suppose Φ_ℓ is the only complex-valued, centered Gaussian Random field on $\mathbb{S}^2$ whose probability distribution is rotation-invariant, and whose samples belong to the irreducible subspace of $\mathbb{L}^2(\mathbb{S}^2)$ spanned by the spherical harmonics $Y_{\ell m}$, $m = -\ell, ..\ell$. Consider a Borel subset $\mathscr{A}$ of $\mathbb{S}^2$, write $|\mathscr{A}|_s$ for its area w.r.t the round metric, and $\mathscr{N}_\mathscr{A}$ for the random variable recording the number of zeroes of $\mathbf{z}$ in $\mathscr{A}$. Set $\Lambda = \frac{2\pi}{\sqrt{\ell(\ell+1)}}$ as above. Then*

$$\frac{\mathbb{E}\left\{\mathscr{N}_\mathscr{A}\right\}}{|\mathscr{A}|_s} = \frac{\pi}{\Lambda^2}.$$

Let us give a sketch of proof of Theorem D: since the only difference with the hyperbolic case is the lack of global coordinates which simplified the presentation there, we think it is better to keep this proof short and refer to our upcoming PhD thesis for full details. A first step is to adapt the formula by Azais and Wschebor (the version on page 22 holds when the field is defined on an open subset of $\mathbb{R}^n$) to prove that $\mathbb{E}\left\{\mathscr{N}_\mathscr{A}\right\} = |\mathscr{A}|_s \frac{V_0}{\pi}$, where V_0 is the variance of any derivative of $\mathfrak{Re}(\mathbf{z})$ at a point p_0 on $\mathbb{S}^2$. Now to evaluate V_0, we use the fact that it is equal to the expectation for the second partial derivative (in any direction) at p_0 of the covariance function Γ of $\mathfrak{Re}(\mathbf{z})$. This expectation does not depend on the chosen direction, and to adapt the arguments in the proof of Theorem A we can use the group-theoretical interpretation of $\Delta_{\mathbb{S}^2}$ as the Casimir operator associated to the action of $SO(3)$ on $\mathbb{S}^2$ (see [40], section 5.7.7). As in the proof of Theorem A, we can then evaluate V_0 as half the value of $\Delta_{\mathbb{S}^2}(\Gamma)$ at p_0, but because $\Delta_{\mathbb{S}^2}\mathbf{z} = \left(\frac{2\pi}{\Lambda}\right)^2 \mathbf{z}$ this half-value turns out to be $\frac{\pi}{\Lambda}^2$, proving Theorem D.

3.2.d An alternative orientation map, with shift-twist symmetry

We have so far been looking for arrangments of $V1$-like receptive profiles on curved (homogeneous) surfaces; for this we used complex-valued random fields. We shall now look for a pinwheel-like structure on the sphere which is of a slightly different kind, perhaps more likely to be of use in discussions which include horizontal connectivity, or which relate to the vestibular system and its interaction with vision. We will also provide a simple criterion on pinwheel densities to distinguish between our two types of spherical maps.
In this subsection $\mathbb{S}^2$ sits as the unit sphere in $\mathbb{R}^3$, and we try to arrange *three*-dimensional abelian Fourier coefficients on the sphere: in other words, we assume each point $\vec{u}$ on $\mathbb{S}^2$ is equipped with a sensor whose receptive profile depends on a plane wave $x \in \mathbb{R}^3 \mapsto exp\,\langle \omega(\vec{u}) \cdot x \rangle$ (this profile could be a three-dimensional Gabor wavelet). Here $\omega(\vec{u}) \in \mathbb{R}^3$ is a linear form on $\mathbb{R}^3$ (so it may be thought of as a vector). Let us assume further that at each point $\vec{u}$, the corresponding sensor neglects everything that happens in directions collinear to $\vec{u}$, so that $\omega(\vec{u}) \cdot v = 0$ as soon as $\vec{v} \perp \vec{u}$.

This kind of arrangment does not seem very interesting if (a part of) the sphere is thought of as a piece of cortical surface, and we do not set it forth as a model for a visual area; yet it would not be completely unreasonable to think of an arrangment like this if $\vec{u}$ were to stand for gaze direction, and it makes sense (not to say that it is useful) to consider a remapping of this structure across the cortical surface (this would displace the interpretation of the pinwheel-like layout, which would only exist at a functional level).

Now, there is a natural operation of the rotation group $SO(3)$ on such arrangments: if R is a rotation and ω is a map as above, then the natural "rotated ω", *viz.*

$$\vec{u} \overset{R_\star\omega}{\longmapsto} R \cdot \omega(R^{-1}\vec{u})$$

is an arrangment of the same kind. Notice that if R is a rotation of axis $\vec{u}$, it shifts the "orientation preference" in $\omega(\vec{u})$.

This formula is familiar from differential geometry; in fact, our set of maps is precisely the set $\Omega^1(\mathbb{S}^2)$ of (vector fields or, more accurately,) *differential 1-forms* on the sphere. Now, let us come back to $\Omega^1(\mathbb{S}^2)$: we can add two such maps, so $\Omega^1(\mathbb{S}^2)$ is a vector space. After a suitable completion, we may consider the Hilbert space $\Omega^1_{\mathbb{L}^2}(\mathbb{S}^2)$ of forms which are square-integrable, and since rotations are unitary maps, and writing $P(g)$ for the map $\omega \mapsto g_\star\omega$ whenever g is a rotation, we get a unitary representation $(\Omega^1_{\mathbb{L}^2}(\mathbb{S}^2), P)$ of the rotation group.

Using this representation may look rather unnatural in biology; but corresponding transformations *have* been discussed in the flat case, though with a very different language: in [8, 6] they are called *Shift-twist transformations*. Indeed, differential forms on $\mathbb{R}^2$ can be identified with functions from $\mathbb{R}^2$ to $\mathbb{C}$, and the natural action on differential forms of a rotation around the origin [15] $(A, 0) \in SE(2)$ is turned in this way into the

15. Because translations have zero derivative, a general element (A, v) of $SE(2)$ then acts on complex-valued functions on $\mathbb{R}^2$ as $f \mapsto \big(x \mapsto Af(-A^{-1}v + A^{-1}x)\big)$.

operation $f \mapsto Af(A^{-1}\cdot)$ on complex-valued functions, which is exactly the shift-twist transformation considered in [8, 6] (compare section 2.3 in [6])

Bringing the horizontal connectivity and notions like the association field into the picture ([38], [10] chap. 4), it seems natural to introduce the (co-)tangent bundle of the surface on which orientation maps are to be developed.

Now, the unitary representation $(\Omega^1_{\mathbb{L}^2}(\mathbb{S}^2), P)$ is of course not irreducible; so in order to get "elementary arrangments", we may look for its irreducible constituents as we did for $\mathbb{L}^2(\mathbb{S}^2)$ and hope that pinwheel-like structures are to be found there.

There is a useful remark here: if f is a real-valued smooth function on the sphere, its derivative df provides us with an element of $\Omega^1_{\mathbb{L}^2}(\mathbb{S}^2)$. What is more, if g is a rotation, then $P(g)df = d\left[\vec{u} \mapsto f(g^{-1}\vec{u})\right]$. So any G-invariant irreducible subspace $\mathcal{H}_\ell$ of $\mathbb{L}^2(\mathbb{S}^2)$ yields a G-invariant irreducible subspace $\mathcal{H}_\ell^{exact}$ of $\Omega^1_{\mathbb{L}^2}(\mathbb{S}^2)$: we need only consider the derivatives of real parts of elements of $\mathcal{H}_\ell$.

All in all, if we start with one of the monochromatic random fields Φ_ℓ, $\ell \geq 1$ and consider the derivative of its real part, we get a random differential form $\varpi_\ell$on $\mathbb{S}^2$ which probes one irreducible factor of $\Omega^1(\mathbb{S}^2)$. What kind of "orientation map" does this correspond to ? Plotting this needs a warning: when ω is a differential form the $\omega(\vec{u})$ appear in different tangent planes as $\vec{u}$ varies, so a picture may be misleading; luckily it is orientation maps we wish to plot, and the projections of the $\omega(\vec{u})$ on a plane through zero give a fine idea of the layout of orientations on each of the hemispheres it cleaves $\mathbb{S}^2$ into. A plot of a projection of ϖ_ℓ^{exact} for $\ell = 10$ is displayed on figure 7.

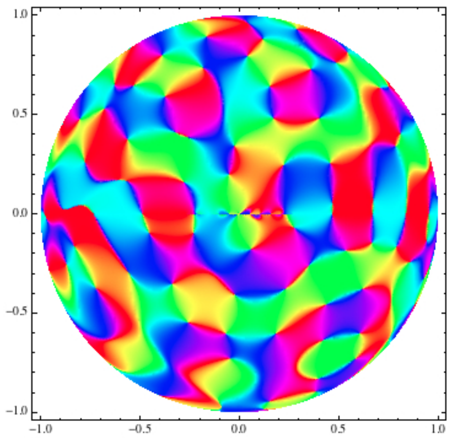


Figure 7: An orientation map on the sphere sampled from a random vector field which has $SO(3)$-shift-twist symmetry. We plotted the restriction to a hemisphere of the random map exploring $\mathcal{H}_{10}^{exact}$; beware that the color coding has a different meaning than in Figs. 1, 2, 5, 6. Here, the sample map is a vector field on the sphere, and there is no complex number; to visualize the direction of the emerging vector at each point, we apply the orthogonal projection from the drawn hemisphere to the "paper" plane, thus getting a vector field on the unit ball of the Euclidean plane, and plot the resulting orientation map using the same color code as in Fig. 1, 2, 5, 6.

Is there more to fields probing other irreducible factors of $\Omega^1_{\mathbb{L}^2}(\mathbb{S}^2)$ than what we see on the $\mathcal{H}_\ell^{exact}$? There is not, for there is a duality operation on $\Omega^1_{\mathbb{L}^2}(\mathbb{S}^2)$ which will allow us to describe all the other irreducible factors. This is the Hodge star: to define it in our very particular case, notice first that if $\vec{u}$ is a unit vector, we get an notion of oriented bases on the plane $\vec{u}^{\perp}$ from “the" usual notion of oriented basis in the ambient space. Then, start with a differential form ω, and shift each of the (co)-vectors $\omega(\vec{u})$ with a rotation of angle $+\frac{\pi}{2}$ in each (co) tangent space; this gives a new form $\star\omega$. Obviously it is orthogonal to ω, and what is more, it commutes with rotations: $g_\star(\star\omega) = \star(g_\star\omega)$ for any rotation g.

Let us write $\mathcal{H}_\ell^{coexact}$ for the image of $\mathcal{H}_\ell^{exact}$ under the Hodge star; since the Hodge star is a G-invariant isometry of $\Omega^1_{\mathbb{L}^2}(\mathbb{S}^2)$, it is a G-invariant irreducible subspace too. Now, using a fundamental result of differential geometry (the Hodge-de Rham theorem), we can deduce from this the decomposition

$$\Omega^1_{\mathbb{L}^2}(\mathbb{S}^2) = \bigoplus_{\ell>0} \left(\mathcal{H}_\ell^{exact} \oplus \mathcal{H}_\ell^{coexact}\right)$$

where the direct sum is orthogonal.

Random differential forms probing the $\mathcal{H}_\ell^{coexact}$ have exactly the same orientation preference layout as those we have already met, except for a difference of "chirality" that corresponds to a global shift of the orientations. We should note here that (the probablility distributions for) our non-twisted fields Φ_ℓ were unchanged under a global shift of the orientations.

While our new maps do resemble the non-twisted orientation maps of the previous paragraph, looking at pinwheel densities will reveal a notable difference. Indeed, although there is a formula of Kac-Rice type for the mean number of critical points of an invariant monochromatic field like $\mathfrak{Re}(\Phi_\ell)$, it involves a Hessian determinant at the place where we earlier met the Jacobian determinant of Φ_ℓ − this was the determinant of a random matrix with independent coefficients, which is not the case for any Hessian (symmetric !) matrix.

We now need to deal with the mean determinant of a random matrix whose coefficients have gaussian distributions but exhibit non-trivial correlations. This seems intractable in full generality; fortunately, our specific spherical problem has been solved recently: in ref. [47], the author proves that the mean total number of critical points of a monochromatic gaussian invariant field, that is, the expectation for the total number of pinwheels (beware this is not a density) in ϖ_ℓ^{exact}, is equivalent to

$$\frac{\ell^2}{\sqrt{3}}$$

as ℓ goes to infinity. Actually, for "finite" ℓ, the total number is given by an explicit but complicated expression.

Note that in our non-twisted, complex-valued random fields Φ_ℓ, the expectation for the total number of pinwheels is equivalent to

$$\ell^2$$

as ℓ goes to infinity. So it is easy, at least in principle, to distinguish the two kinds of orientation maps: one needs only a single quantitative measurement.

4 Discussion

In this paper, we started from a reformulation of existing work by Wolf, Geisel and colleagues, with the aim to understand the crucial symmetry arguments used in models with the help of noncommutative harmonic analysis, which is often a very well-suited tool for using symmetry arguments in analysis and probability. Understanding these Euclidean symmetry arguments from a conceptual standpoint showed us that Euclidean geometry at the cortical level is a way to enforce conditions that are not specific to Euclidean geometry but have a meaning on every “symmetric enough" space, and we thus saw how a unique Gaussian random field providing V1-like maps can be associated to each irreducible “factor" in the Plancherel decomposition of the Hilbert space of square-integrable functions on the Euclidean plane, the hyperbolic plane and the sphere. We proved that in these three settings, when scaled with the typical value of column spacing, monochromatic invariant fields exhibit a pinwheel density of π. Theorems A’ and D’ in the appendix prove that the same result holds when the monochromaticity condition is dropped: in other words, a pinwheel density of π appears as a signature of (shift) symmetry. Since pinwheel densities can be measured in individual sample maps thanks to the ergodicity properties of invariant gaussian fields (see [30], section 6.5), this yields a criterion to see whether an individual map (which can not be itself invariant !) is likely to be a sample from a field with an invariance property, whether the map be drawn on a flat region or on a curved, homogeneous enough region. In the spherical case, also we saw that the number of pinwheels in the map can in principle distinguish between rotation-invariance and shift-twist symmetry; to see whether this observation can be turned into a precise criterion distinguishing the various kinds of invariance from actual measurements on individual sample maps, it would probably be interesing to see whether there is anything to be said of pinwheel densities in Euclidean or hyperbolic maps with shift-twist symmetry, and of the mean column spacing in shift-twist symmetric maps.

Since our aim was to understand the role of symmetry arguments, one aspect restricting the scope of our constructions in a fundamental way is our focusing on homogeneous spaces rather than spaces with variable curvature. Of course we have good technical reasons for this: the way symmetry arguments are used in existing discussions made it natural to focus on those two-dimensional spaces which have a large enough symmetry group, and our constructions are entirely based on exploiting the presence of this symmetry group. One might wish to make the setting less restrictive, especially since the places where the surface of real brains is closest to a homogeneous space are likely to be the flat parts. But using analogues of symmetry arguments on nonsymmetric spaces is a major challenge in (quantum) field theory, and if one wished to start from the reformulation we gave of Wolf and Geisel’s work in section 2.4, generalizing the arguments of this paper to find V1-like maps on riemannian manifolds on nonconstant curvature would be formally analogous to adapting Wigner’s description of elementary particles on Minkowski spacetime to a general curved spacetime − a challenge indeed ! Answering this challenge would bring us close to the two-dimensional models from quantum field theory or statistical mechanics, and make us jump to infinite-dimensional “phase spaces" (and would-be groups). This is a step the author is not ready to take, and it is likely that simpler ways to study the nonhomogeneous case would come with shifting the focus from mature maps back to development models.

Indeed, readers familiar with development models have perhaps been puzzled by another aspect of our paper, which is the fact that we used Gaussian Random Fields as the setting for our constructions: Gaussian fields provide sample maps which look very much like orientation maps, and as we emphasized the statistical properties of their zero set are very strikingly reminiscent of what is to be found in real maps, but there are appreciable and measurable differences between the output of invariant gaussian fields and real orientation maps (see for instance a discussion in [73]). As we recalled in the Introduction, it is likely that Gaussian fields provide a better description for the early stage of cortical map development, but that the Gaussian description later acquires drawbacks because it is not compatible with the nonlinearities essential to realistic development scenarii.

To our knowledge, many of the most successful models for describing the mature stage of orientation preference maps are variations on the long-range interaction model of Wolf, Kaschube et al [68, 69, 1], the mature map $\mathbf{z}$ evolves from an undetermined (random) initial stage (not assumed to be Gaussian) through

$$\partial_t \mathbf{z} = L_\Lambda(\mathbf{z}) + N_{\gamma,\sigma}(\mathbf{z})$$

Here L_Λ is a Swift-Hohenberg operator

$$\mathbf{z} \mapsto r\mathbf{z} - ((\frac{2\pi}{\Lambda})^2 + \nabla^2)^2 \mathbf{z},$$

γ is a real number between zero and two, σ is a positive number and $N_{\gamma,\sigma}$ is the following nonlinear operator:

$$N[\mathbf{z}] := x \mapsto (1-\gamma)|\mathbf{z}(x)|^2\mathbf{z}(x) - (2-\gamma)\int_{\mathbb{R}^2} K_\sigma(x-y)\left(\mathbf{z}(x)|\mathbf{z}(y)|^2 + \frac{1}{2}\bar{\mathbf{z}}(x)\mathbf{z}(y)^2\right) dy.$$

Allowing $\mathbf{z}$ to evolve from an initial fluctuation, when $\gamma < 1$ and when σ/Λ is large enough equation (1) leads first to an invariant, approximately Gaussian field (thanks to an application of the central limit theorem to a linearized version of Eq. (1), see [5]), then to non-Gaussian quasiperiodic V1-like random fields.

This interaction model can be easily adapted to define a non-linear partial differential equation on any riemannian manifold: a riemannian metric, say on M, comes with a natural laplacian Δ_M and a volume form $dVol_M$, so we can define L_Λ^M as

$$\mathbf{z} \mapsto r\mathbf{z} - ((\frac{2\pi}{\Lambda})^2 + \Delta_M)^2 \mathbf{z},$$

and use the geodesic distance $d_M(x,y)$ between any two points x, y of M to define

$$N_\gamma^M[\mathbf{z}] := x \mapsto (1-\gamma)|\mathbf{z}(x)|^2\mathbf{z}(x) - (2-\gamma)\int_M e^{-\frac{d_M(x,y)^2}{2\sigma^2}}\left(\mathbf{z}(x)|\mathbf{z}(y)|^2 + \frac{1}{2}\bar{\mathbf{z}}(x)\mathbf{z}(y)^2\right) dVol_M(y).$$

A non-Euclidean version of the long-range interaction model on M would then simply be

$$\partial_t \mathbf{z} = L_\Lambda^M(\mathbf{z}) + N_{\gamma,\sigma}^M(\mathbf{z}). \tag{4.1}$$

As Wolf and coworkers point out (see for instance the supplementary material in [1], section 2), this partial differential equation is the Euler-Lagrange equation of a variational problem, so solutions are guaranteed to converge to stable stationary states as time wears on. Wolf and colleagues showed (using numerical studies) that when σ/Λ is large enough, V1-like maps are among the stable solutions in the Euclidean case. On arbitrary riemannian manifolds, however, there is no way to guarantee that structure-rich stable solutions of the above PDE exist; it would certainly be worth examining, at least with numerical simulations, but this is beyond the author's strengths at present. It is perhaps natural to imagine that given the analogy between maps obtained by truncation from invariant GRFs and the output of the long-range interaction model, the constructions in this paper are a strong indication that on symmetric spaces, the stable solutions of (1) include maps which look like those of figs. 5-7, and that the difference between those and the monochromatic invariant fields studied in this paper is analogous to the difference between experimental maps, or at least the output of (1) in the Euclidean case, and maps sampled from invariant Gaussian fields on $\mathbb{R}^2$.

In adult animals measurements seem to indicate that the structure of mature maps departs from that of maps sampled from invariant Gaussian fields; remarkably, there is experimental evidence for the fact that a pinwheel density of π, which in a Gaussian initial stage appears as a signature of Euclidean symmetry as we saw, is maintained during development in spite of the important refinments in cortical circuitry and the departure from Gaussianity that they induce [1]. Independently of modelling details, we see that geometrical invariance can be measured in principle, even on individual maps: upon evaluating local column spacings (with respect to geodesic length in the curved case) and performing space averaging, one gets a length scale Λ; when scaling pinwheel density with respect to Λ, observing a value of π is a strong indication that geometrical invariance on the cortical surface is an important ingredient in development.

In addition to this, one might think that arranging neurons and their receptive profiles on a homogeneous enough space has benefits from the point of view of information processing: by allowing the conditions of homogeneity and isotropy to make sense, a constant curvature could help distribute the information about the stimulus in a more uniform way (note that as the eyes move constantly, a given image is processed by many different areas in V1 in a relatively short time). Neurons receiving inputs from several adjacent regions of V1 could then handle spike statistics which vary little as the sensors move, and have a more stable worldview.

Competing Interests.

The author declares that he has no competing interests.

Bibliography

[1] Kaschube, M., Schnabel, M., Lowel, S., Coppola, D. M., White, L. E., & Wolf, F: *Universality in the evolution of orientation columns in the visual cortex.* Science 2010, **330(6007)**:1113-1116.

[2] Keil, W., Kaschube, M., Schnabel, M., Kisvarday, Z. F., Löwel, S., Coppola, D. M., & Wolf, F.: *Response to Comment on "Universality in the Evolution of Orientation Columns in the Visual Cortex ".* Science 2012, **336(6080)**:413-413.

[3] Wolf, F., & Geisel, T.: *Spontaneous pinwheel annihilation during visual development.* Nature 1998, **395(6697)**: 73-78.

[4] Kaschube, M., Wolf, F., Geisel, T., & Lowel, S.: *Genetic influence on quantitative features of neocortical architecture..* The Journal of Neuroscience 2002, **22(16)**:7206-7217.

[5] Wolf, F., & Geisel, T.: *Universality in visual cortical pattern formation. .* Journal of Physiology-Paris 2003, **97(2)**: 253-264.

[6] Schnabel, M., Kaschube, M., Löwel, S., & Wolf, F.: *Random waves in the brain: Symmetries and defect generation in the visual cortex. .* The European Physical Journal Special Topics 2007, **145(1)**: 137-157.

[7] Miller, K.D. $\pi =$ *Visual Cortex.* Science 2010 **330**:1059-1060.

[8] Schnabel M: *A Symmetry of the Visual World in the Architecture of the Visual Cortex.* PhD thesis. University of Goettingen; 2008.

[9] Reichl L, Heide D, Lowel S, Crowley JC, Kaschube M, et al. *Coordinated Optimization of Visual Cortical Maps (I) Symmetry-based Analysis.* PLoS Comput Biol 2012 **8(11)**

[10] Petitot, J. *Neurogéométrie de la vision: modeles mathematiques et physiques des architectures fonctionnelles.* Paris: Editions Ecole Polytechnique; 2008.

[11] Hubel, D. H., & Wiesel, T. N. *Receptive fields, binocular interaction and functional architecture in the cat's visual cortex.* The Journal of physiology 1962, **160(1)**, 106.

[12] Hubel, D. H., & Wiesel, T. N. *Receptive fields of single neurones in the cat's striate cortex.* The Journal of physiology 1959 **148(3)**:574-591, 1959.

[13] Hubel, D. H., & Wiesel, T. N. *Receptive fields and functional architecture of monkey striate cortex.* The Journal of physiology 1968, **195(1)**, 215-243.

[14] Hubel, D. H.. *Eye, brain, and vision.* Scientific American Library:Scientific American Books; 1995.

[15] Ohki, K., Chung, S., Kara, P., Hubener, M., Bonhoeffer, T., & Reid, R. C. *Highly ordered arrangement of single neurons in orientation pinwheels.* Nature 2006, **442(7105)**, 925-928.

[16] Yu H, Farley B J, Jin D Z, & Sur M. *The coordinated mapping of visual space and response features in visual cortex.* Neuron 1995, **47(2)**:267–280.

[17] Swindale, N.V. *A model for the formation of orientation columns.* Proc. Royal Society of London. Series B: Biological Sciences 1982 **215 (1199)**, 211-230.

[18] Swindale, N. V. *The development of topography in the visual cortex: A review of models.* Network: Computation in neural systems 1996, **7(2)**, 161-247.

[19] Niebur, E., Worgotter, F., *Design principles of columnar organization in visual cortex.* Neural Computation 1994 **6 (4)**, 602-614.

[20] Bonhoeffer T, Grinvald A. *The layout of iso-orientation domains in area 18 of cat visual cortex: optical imaging reveals a pinwheel-like organization* J. Neurosci 1996. **13**, 4157

[21] Bonhoeffer and Grinvald A. *Iso-orientation domains in cat visual cortex are arranged in pinwheel-like patterns.* Nature 1991, **353(6343)**:429-431.

[22] Barbieri, D., Citti, G., Sanguinetti, G., & Sarti, A. *An uncertainty principle underlying the functional architecture of V1.* Journal of Physiology-Paris 2012, **106(5)**, 183-193.

[23] Dahlem, M. A., & Tusch, J. *Predicted selective increase of cortical magnification due to cortical folding.* The Journal of Mathematical Neuroscience (JMN) 2012, **2(1)**, 1-26.

[24] Berry, M. V.; Dennis, M. R. *Phase singularities in isotropic random waves.* Royal Society of London Proceedings Series A 2000, **456(2001)**: 2059

[25] Azais J M, Wschebor M, *Level Sets and extrema of random processes and fields.* New York: John Wiley and Sons; 2009.

[26] Van Hooser S D, Heimel J A, Nelson S B. *Functional cell classes and functional architecture in the early visual system of a highly visual rodent.* Progress in brain research 2005, **149**:127-145.

[27] Kaschube, M. *Neural maps versus salt-and-pepper organization in visual cortex.* Current opinion in neurobiology 2014, **24**: 95-102.

[28] Abrahamsen, P. *A review of Gaussian random fields and correlation functions.* Norsk Regnesentral, 1995.

[29] Azais, J. M., Leon, J. R., Wschebor, M. *Rice formulae and Gaussian waves.* Bernoulli 2011 **17(1)**, 170-193.

[30] Adler, R. J., & Taylor, J. E. *Random fields and geometry* Berlin: Springer; 2009.

[31] Cramer, H. and Leadbetter, M.R. *Stationary and related stochastic processes. Sample function properties and their applications.* New York: John Wiley and Sons, 1967

[32] Bosking W H, Zhang Y, Schofield B, and Fitzpatrick D. *Orientation selectivity and the arrangement of horizontal connections in tree shrew striate cortex.* J. Neurosci. 1997, **17(6)**:2112-2127.

[33] Hepp K. *On Listing's Law.* Communications in Mathematical Physics 1990, **132**, pp. 285-292.

[34] Yaglom, A. M. *Second-order homogeneous random fields.* In Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, Volume 2: Contributions to Probability Theory. The Regents of the University of California, 1961.

[35] Chalupa, L. M., & Werner, J. S. (Eds.). (2004). The visual neurosciences (Vol. 1). Cambridge, MA: MIT press.

[36] Ernst, U. A., Pawelzik, K. R., Sahar-Pikielny, C., & Tsodyks, M. V. *Intracortical origin of visual maps.* Nature neuroscience 2001, **4(4)**, 431-436.

[37] Maffei, L., & Galli-Resta, L. *Correlation in the discharges of neighboring rat retinal ganglion cells during prenatal life.* Proceedings of the National Academy of Sciences 1990, **87(7)**, 2861-2864.

[38] Bressloff, P. C., Cowan, J. D., Golubitsky, M., Thomas, P. J., & Wiener, M. C. *Geometric visual hallucinations, Euclidean symmetry and the functional architecture of striate cortex.* Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences 2001, **356(1407)**, 299-330.

[39] Chossat, P., & Faugeras, O. *Hyperbolic planforms in relation to visual edges and textures perception.* PLoS computational biology 2009, **5(12)**, e1000625.

[40] Gurarie, D. *Symmetries and Laplacians: Introduction to Harmonic Analysis, Group Representations and Applications.* London: Courier Dover Publications, 2007.

[41] Vilenkin, N. I. *Representation of Lie Groups and Special Functions: Simplest Lie Groups, Special Functions, and Integral Transforms* Berlin:Springer, 1991.

[42] Wigner E. *On unitary representations of the inhomogeneous Lorentz group.* Annals of mathematics 1939 **40(1)**, 149-204.

[43] Weyl, H. *The classical groups: their invariants and representations.* Princeton university press, 1946.

[44] Helgason, S. *Differential geometry, Lie groups, and symmetric spaces* New York: Academic press; 1979.

[45] Helgason, S. *Groups & Geometric Analysis: Radon Transforms, Invariant Differential Operators and Spherical Functions* New York: Academic Press; 1984

[46] Helgason, S. *Geometric analysis on symmetric spaces* American Mathematical Society, 2008.

[47] Nicolaescu, L. *Critical sets of random smooth functions on compact manifolds.* arXiv preprint 2010: arXiv:1008.5085

[48] Sugiura, M. *Unitary representations and harmonic analysis: an introduction.* New York: Elsevier, 1990.

[49] Gelbart, S. *An elementary introduction to the Langlands program.* Bulletin of the American Mathematical Society 1990, *10(2)*: 177-219.

[50] Felix Klein. *A comparative review of recent researches in geometry, translated by M. W. Haskell*, Bull. New York Math. Soc. 1892 **2**, 215-249.

[51] Harish-Chandra, *Spherical functions on a semisimple Lie group. I*, American Journal of Mathematics 1958 **80**: 241–310

[52] Cover, T. M., & Thomas, J. A. *Elements of information theory.* New York: John Wiley & Sons, 1991.

[53] Knapp, A. W, *Lie groups: beyond an introduction* Berlin: Springer, 2008.

[54] Mackey, G. W. *Harmonic analysis as the exploitation of symmetry-a historical survey.* Bulletin of the American Mathematical Society 1980 **3(1)**, 543-698.

[55] Weyl, H, *Gruppentheorie und quantenmechanik.* Zurich, 1928.

[56] Sharpe, R. W. *Differential geometry: Cartan's generalization of Klein's Erlangen program* Berlin: Springer, 1997.

[57] G. C. De Angelis, I. Ohzawa, and R. D. Freeman, Spatiotemporal organization of simple-cell receptive fields in the cat's striate cortex. I. General characteristics and postnatal development. J. Neurophysiol. 69, 1091-1117 (1993)

[58] Cocci G, Barbieri D. & Sarti A. *Spatiotemporal receptive fields of cells in V1 are optimally shaped for stimulus velocity estimation.* Journal of the Optical Society of America A (2001): **29(1)**: 130-138.

[59] Escher, M.C. *L'oeuvre graphique.* Cologne: Taschen, 2001.

[60] Dudley, R. M. (2002). *Real analysis and probability.* Cambridge University Press (Vol. 74).

[61] Karman, G. P., Beijersbergen, M. W., van Duijl, A. & Woerdman, J. P., *Creation and annihilation of phase singularities in a focal field.* Optics Letters 1996 **22**, 1503-1505.

[62] Nye, J F, & Berry, M V, *Dislocations in wave trains*, Proc. R. Soc. Series A (1974), **336**: 165-90

[63] Inouye, S., Gupta, S., Rosenband, T., Chikkatur, A. P., Gorlitz, A., Gustavson, T. L., ... & Ketterle, W. *Observation of vortex phase singularities in Bose-Einstein condensates*, Physical review letters (2001) **87(8)**, 080402.

[64] Barbieri D. & Citti, G. *Reproducing kernel Hilbert spaces of CR functions for the Euclidean motion group.* Analysis and applications (2015) **13** 331.

[65] Nauhaus I, Nielsen K, Disney A, Callaway E. *Orthogonal micro-organization of orientation and spatial frequency in primate primary visual cortex.* Nature Neuroscience 2012 **15**, 1683-1690

[66] Levy P. *Processus stochastiques et mouvement brownien.* Paris: Gauthiers Villars, 1948.

[67] Courant R. & Hilbert D. *Methods of Mathematical Physics.* New York: Interscience, 1953.

[68] Wolf, F. *Symmetry, Multistability, and Long-Range Interactions in Brain Development.* Physical Review Letters 2005, **95**:208701

[69] M. Kaschube, M. Schnabel, F. Wolf. *Self-Organization and the Selection of Pinwheel Density in Visual Cortical Development.* New Journal of Physics (2008) **10**: 015009.

[70] Rajimehr, R. & Tootell, R.B.H. *Does Retinotopy Influence Cortical Folding in Primate Visual Cortex?* The Journal of Neuroscience (2009), **29(36)**:11149 –11152.

[71] Kuo, H.H. *Introduction to Stochastic Integration.* Berlin: Springer, 2006.

[72] Van Essen, D.C. (2006) *Cerebral cortical folding patterns in primates: Why they vary and what they signify.* In: Evolution of Nervous Systems (ed. J.H. Kaas); Elsevier Press, San Diego, pp. 267-276.

[73] Afgoustidis, A. *Monochromaticity of orientation maps in V1 implies minimum variance for hypercolumn size.* The Journal of Mathematical Neuroscience (2015), **5**:10.

Chapter 4

Invariant Gaussian Fields on Homogeneous Spaces : Explicit Constructions and Geometric Measure of the Zero-Set

Contents

Abstract. This chapter, a mathematical sequel to Chapter 3, turns to invariant Gaussian random fields defined on a homogeneous space of *any* dimension. I first indicate, building on early results on Yaglom, how the available information on representation-theory-related special functions makes it possible to give completely explicit descriptions of these fields in many cases of interest. I then turn to the expected size of the zero-set for these fields: extending the two-dimensional results above, I show that each invariant field comes with a natural unit of volume (defined in terms of the geometrical redundancies in the field) with respect to which the average size of the zero-set depends only on the dimension of the source and target spaces, and not on the precise symmetry exhibited by the field.

1 Introduction

Interest for Gaussian random fields with symmetry properties has risen recently. While it is not surprising that these fields should have many applications (Kolmogorov insisted as early as 1944 that they should be relevant to mathematical discussions of turbulence), a short list of recent domains in which they appeared will help me describe the motivation for this paper.

• Optics and the Earth sciences. Suppose a wave is emitted at some point, but thereafter undergoes multiple diffractions within a disordered material (like a tainted glass, or the inside of the Earth). If the material is disordered enough, beyond it one will observe a superposition of waves propagating in somewhat random directions, with somewhat random amplitudes and phases. In Sismology and in Optics, there are theoretical and practical benefits in treating the output as a single realization of a complex-valued gaussian field, which inherits symmetry properties from those of the material which diffracted the waves. See [10] for Optics, [37] for Sismology.

• Astrophysics. The Cosmic Microwave Background (CMB) radiation is understood to be an observable relic of the “Big Bang". There are fluctuations within it [1]: the frequency of the radiation changes very slightly around the mean value as one looks to different parts of the sky. The fact that the variations are small is essential for cosmology of course, but their precise structure is quite as important: they are supposed to be a relic of the slight homogeneities which made it possible for the galaxies and stars to take shape. Because it is customary to assume that the universe is, and has always been, isotropic on a large scale, a much-discussed model for the CMB treats it as a single realization of an isotropic gaussian field on the sphere [2]. This has of course prompted mathematical developments [21], as well as motivation for a recent monograph by A. Malyarenko [20] on random fields with symmetry properties.

• Texture modelling and synthesis. Many homogeneous-looking regions in natural images, usually called *textures*, are difficult to distinguish with the naked eye from realizations of an appropriate stationary Gaussian random field. Thus, stationary Gaussian fields are a simple and natural tool for image synthesis; an appreciable advantage of this is that thanks to the ergodicity properties of Gaussian fields ([2], chapter 6), the probability distribution of a stationary Gaussian random field can be roughly recovered from a single realization: measuring correlations in a single sample yields a good approximation for the covariance function of the field, and one can draw new examples of the given texture from it: this is widely used in practice. See [31, 27, 14]. Textures obviously have a meaning on homogeneous spaces as well as Euclidean space [3], and the mathematical generalization of widely-used image-processing tools to curved spaces should feature homogeneous fields [4].

1. These were detected in 1992; see G. Smoot and J. Mather’s Nobel lectures [28, 23]

2. In fact, the expectation function of the field should not be a constant, because the CMB has rather large-scale fluctuations, famously attributed to a Doppler effect due to the metric expansion of space, in addition to the above-mentioned variations. So it is the centered version (with the large-scale fluctuations substracted off) that should be isotropic. Instead of being Gaussian, the field could also be a function of an underlying Gaussian field: this seems to be a prevailing hypothesis; see [21].

3. The example images in [14] can obviously imagined on the surface of a sphere or in a hyperbolic plane.

4. This does not mean that it will be faster to work with on a computer than a version less naturally suited to the curvature of the manifold!

• Neuroscience (more detailed discussions can be found in [3, 4]). In the primary visual cortex of mammals, neurons record several local features of the visual input, and the electrical activity of a given neuron famously depends on the presence, in its favourite region of space, of oriented stimuli (“edges"): there are neurons which activate strongly in the presence of vertical edges in the image, others which react to oblique edges, and so on. The map which, to a point of the cortical surface, assigns the "orientation preference" of the neuron situated there (the stimulus direction which maximizes the electrical activity of the neuron) has been observed to be continuous in almost all (though not all) mammals, and to have strikingly constant geometrical properties across species and individuals. Prevailing models for the early stage in the development of these cortical maps treat the arrangment in a given individual as a single realization of a Gaussian random field on the cortical surface, with the orientation map obtained after taking the argument; a key to the success of the models is the assumption, meant to reflect the initial homogeneity of the biological tissue, that when the cortical surface is identified with a Euclidean plane, the underlying Gaussian field is homogeneous and isotropic.

In at least two of these fields, Optics and Neuroscience, the *zeroes* of stationary gaussian fields have attracted detailed attention. In a heated debate on the evolution of the early visual system in mammals (see [24]), the mean number of zeroes in a region with a given area has been taken up as a criterion to decide between two classes of biological explanations for the geometry of cortical maps. Experiments show the mean value to be remarkably close to π with respect to an appropriate unit of area (re-defined for Gaussian fields in section 5 below). This strikingly coincidates with the exact mean value obtained for stationary isotropic Gaussian fields by Wolf and Geisel in related work, and independently by Berry and Dennis in optics-related work (there the zeroes are points where the light goes off, or the sound waves cancel each other : Berry and Dennis call them "lines of darkness, or threads of silence"). Along with the key role symmetry arguments play in the discussion of the visual cortex, the remarkable coincidence is one of my motivations for generalizing to arbitrary homogeneous spaces the Euclidean-and-planar results which appeared in Optics and Neuroscience.

These recent developments take up an old theme: understanding the properties of the level sets of (the paths of) a random field is a classical subject in the theory of stochastic processes.

This paper is a mathematical follow-up on [3] ; some of the results below have been announced (with incorrect statements !) in the appendix to that article. It has two relatively independent aims:

• Describe invariant Gaussian fields on homogeneous spaces as explicitly as possible,
• Study the mean number of zeroes, or the average size of the zero-set, of an invariant field in a given region of a homogeneous space.

The first problem has been solved in the abstract by Yaglom in 1961 [36] using the observation (to be recalled in section 2.1) that the possible correlation functions of homogeneous complex-valued fields form a class which has been much studied in the representation theory of Lie groups. Section 2.2 is a summary (with independent proofs) of the consequences of his results that I will use. Since Yaglom’s time, representation theory has grown to incorporate several more concrete constructions, and in section 3 below, I show that explicit descriptions (that can be worked with on a computer) are possible on many spaces of interest, including symmetric spaces. Section 3.1 also includes simple facts which show that on a given manifold, not all transitive Lie groups can give rise to invariant random fields with continuous trajectories.

Turning to the second problem, what I show below is that when expressed in a unit of volume appropriate to the field (defined in section 4 for real-valued fields and at the beginning of section 5 for others), the average size of the zero-set does not depend on the group acting, but only on the dimension of the homogeneous space on which the field is defined and that of the space in which it takes its values. When looking at a single realization of a random field, observing the average size for the zero-set expressed by Theorem 2 below can be viewed a signature that the field has a symmetry, regardless of the fine structure of the symmetry involved.

2 Invariant real-valued gaussian fields on homogeneous spaces

2.1 Gaussian fields and their correlation functions

Suppose X is a smooth manifold and V a finite-dimensional Euclidean space. A *Gaussian field on X with values in V* is a random field $\mathbf{\Phi}$ on X such that for each n in $\mathbb{N}$ and every n-tuple $(x_1, ...x_n)$ in X^n, the random vector $(\mathbf{\Phi}(x_1), ...\mathbf{\Phi}(x_n))$ in V^n is a Gaussian vector. A Gaussian field $\mathbf{\Phi}$ is *centered* when the map $x \mapsto \mathbb{E}[\mathbf{\Phi}(x)]$ is identically zero, and it is *continuous*, resp. *smooth*, when $x \mapsto \mathbf{\Phi}(x)$ is almost surely continuous, resp. smooth.

In this paper, our space X will be a smooth manifold equipped with a smooth and transitive action $(g, x) \mapsto g \cdot x$ of a Lie group G. Choose x_0 in X and write K for the stabilizer of x_0 in G. A Gaussian field on X with values in V is *invariant* when the probability distribution of $\mathbf{\Phi}$ and that of the Gaussian field $\mathbf{\Phi} \circ (x \mapsto g \cdot x)$ are the same for every g in G.

The case in which V equals $\mathbb{R}$ is of course important. If Φ is a real-valued Gaussian field on X, its *covariance function* is the (deterministic) map $(x, y) \mapsto \mathbb{E}[\Phi(x)\Phi(y)]$ from $X \times X$ to $\mathbb{R}$. A real-valued Gaussian field is *standard* if it is centered and if $\Phi(x)$ has unit variance at each $x \in X$.

When describing scalar-valued Gaussian fields with symmetry properties, we shall see in the next subsection that the relationship with representation theory makes it useful that the covariance function, and thus the field as well, be allowed to be complex-valued rather than real-valued. A precise word about the kind of complex-valued Gaussian fields we need is perhaps in order here.

A *circularly symmetric Gaussian variable* is a complex-valued random variable whose real and imaginary parts are independent, identically distributed real Gaussian variables. A *circularly symmetric complex Gaussian field on X* is a Gaussian random field Z on X with values in the vector space $\mathbb{C}$, with the additional requirement that $(x, y) \mapsto \mathbb{E}[Z(x)Z(y)]$ be identically zero. Note that while this imposes that $Z(x)$ be circularly symmetric for all x, this does not necessitate that $\mathfrak{Re}(Z)(x)$ and $\mathfrak{Im}(Z)(y)$ be uncorrelated if x is not equal to y.

The *correlation, or covariance, function* of a circularly symmetric complex Gaussian field on X is the (deterministic) map $(x, y) \mapsto \mathbb{E}[Z(x)Z(y)^\star]$ from $X \times X$ to $\mathbb{C}$, where the star denotes complex conjugation. A *standard complex Gaussian field on* X is a circularly symmetric complex Gaussian field on X such that $\mathbb{E}[Z(x)Z(x)^\star] = 1$ for all x.

Note that the real part of the covariance function of a circularly symmetric complex-valued Gaussian field is twice the covariance function of the real-valued Gaussian field obtained by considering its real part. A circularly symmetric complex Gaussian field on X has a real-valued correlation function if and only if its real and imaginary parts are independent *as processes.*

We are now ready for the classical theorem which describes the correlation functions of standard complex Gaussian fields, those of real-valued Gaussian fields being a particular case as we saw (see however [1] for a separate description of the real case).

Proposition 2.1 (see for instance [18], section 2.3)**.** *Suppose C is a deterministic map from $X \times X$ to $\mathbb{C}$. Then it is the covariance function of a continuous (resp. smooth), invariant, standard complex-valued Gaussian field if and only if it has the following properties.*

(a) The map C is continuous (resp. smooth);

(b) for each x, y in X and every g in G, $C(gx, gy) = C(x, y)$;

(c) for every x in X, $C(x, x) = 1$;

(d) (positive-definiteness) for each n in $\mathbb{N}$ and every n-tuple $(x_1, ...x_n)$ in X^n, the hermitian matrix $(C(x_i, x_j))_{1 \leq i,j \leq n}$ is positive-definite.

If $\mathbf{\Phi}_1$ and $\mathbf{\Phi}_2$ are continuous (resp. smooth), invariant, standard complex-valued Gaussian fields with covariance function C, then they have the same probability distribution.

A consequence is that there is a left-and-right K-invariant continuous (resp. smooth) function Γ on G, taking the value one at 1_G, such that $C(gx, x) = \Gamma(g)$ for every g in G and every x in X. Proposition 2.1 thus says that taking covariance functions yields a natural bijection between

• probability distributions of continuous (resp. smooth), invariant, standard complex-valued Gaussian fields on $X = G/K$,
and
• positive-definite, continuous (resp. smooth), K-bi-invariant functions on G, taking the value one at 1_G.

A positive-definite, continuous, complex-valued function on G which takes the value one at 1_G is usually called a *state* of G. We are thus looking for the K-bi-invariant (and smooth, if need be,) states of G.

2.2 How invariant Gaussian fields correspond to group representations

This subsection describes some results due to Yaglom [36], although the presentation differs slightly because I would like to give direct proofs.

Unitary representations of G are a natural source of positive-definite functions: if U is a continuous morphism from G to the unitary group $U(\mathcal{H})$ of a Hilbert space $\mathcal{H}$, then for every unit vector v in $\mathcal{H}$, $g \mapsto \langle v, U(g)v \rangle$ is a state of G. In fact if m is a state of G, there famously is [5] a Hilbert space $\mathcal{H}_m$ and a continuous morphism from G to $U(\mathcal{H}_m)$, as well as a unit vector v_m in $\mathcal{H}_m$, such that [6] $m(g) = \langle v_m, U(g)v_m \rangle$.

The study of K-bi-invariant states is a classical subject when (G, K) is a *Gelfand pair*, that is, when G is connected, K is compact and the convolution algebra of K-bi-invariant integrable functions on G is commutative [7].

It is immediate from the definition that a state of G is a bounded function on G ; thus the K-bi-invariant states of G form a convex subset $\mathcal{C}$ of the vector space $L^\infty(G)$ of bounded functions. Viewing $L^\infty(G)$ as the dual of the space $L^1(G)$ of integrable functions (here we assume a Haar measure is fixed on the − automatically unimodular − group G), and equipping it with the weak topology, $\mathcal{C}$ appears as a relatively compact, convex subset of $L^\infty(G)$ because of Alaoglu's theorem.

The extreme points of $\mathcal{C}$ are usually known as *elementary spherical functions for the pair* (G, K). Their significance to representation theory is that they correspond to *irreducible* unitary representations: if m is a state of G and $(\mathcal{H}, U, v)$ is such that $m = g \mapsto \langle v, U(g)v \rangle$ as above, then m is an elementary spherical function for (G, K) if and only if the unitary representation U of G on $\mathcal{H}$ irreducible [8]. The condition of K-bi-invariance translates into the existence of a K-fixed vector in $\mathcal{H}$.

When (G, K) is a Gelfand pair, the unitary irreducible representations of G which have a K-fixed vector have the subspace of K-fixed vector one-dimensional and not larger: a consequence is that different elementary spherical functions correspond to nonequivalent class-one [9] representations of G. So the Gelfand-Naimark-Segal construction yields a bijection between elementary spherical functions for (G, K) and unitary irreducible representations of G having a K-fixed vector.

To come back to the description of general K-bi-invariant states in terms of the extreme points of $\mathcal{C}$, the Choquet-Bishop-de Leeuw representation theorem (a measure-flavoured generalization of the Krein-Milman theorem) exhibits a general K-bi-invariant state as a "direct integral" of elementary spherical functions, in a way that mirrors the (initially more abstract) decomposition of the corresponding representation of G into irreducibles.

---

5. This is the Gelfand-Naimark-Segal construction: on the vector space $\mathcal{C}_c(G)$ of continuous, compactly-supported functions on a locally compact second countable unimodular group G, we can consider the bilinear form $\langle f, g \rangle_m := \int_{G^2} m(x^{-1}y) f(x) \bar{f}(y) dx dy$. It defines a scalar product on $\mathcal{C}_c(G)/\{f \in \mathcal{C}_c(G)\ ,\ \langle f, f \rangle_m = 0\}$, and we can complete this into a Hilbert space $\mathcal{H}_m$; the natural action of G on $\mathcal{C}_c(G)$ yields a unitary representation of G on $\mathcal{H}_m$.

6. The linear functional $f \mapsto \int_G f \bar{m}$ extends to a bounded linear functional on $\mathcal{H}_m$, and the Riesz representation theorem yields one v_m in $\mathcal{H}_m$ which has the desired property.

7. The subject of positive-definite functions becomes tractable because the Gelfand spectrum of this commutative algebra furnishes a handle on positive-definite functions through the elementary spherical functions to be defined just below.

8. Indeed, should there be $U(G)$-invariant subspaces $\mathcal{H}_1, \mathcal{H}_2$ such that $\mathcal{H} = \mathcal{H}_1 \oplus \mathcal{H}_2$, orthogonal direct sum, writing $v = v_1 + v_2$ with v_i in $\mathcal{H}_i$, one would have $m(g) = \langle v_1, U(g)v_1 \rangle + \langle v_2, U(g)v_2 \rangle$, and $g \mapsto \langle \frac{v_i}{\sqrt{2}}, U(g) \frac{v_i}{\sqrt{2}} \rangle$ would be a positive-definite function, thus m would not be an extreme point of $\mathcal{C}$. The reverse implication is just as easy using the Gelfand-Naimark-Segal construction.

9. A class-one representation is an irreducible representation which has nonzero K-fixed vectors

To be precise, let $\mathbf{\Lambda}$ be the space of extreme points of $\mathcal{C}$, a topological space if one lets it inherit the weak topology from $L^\infty(G)$. Then Choquet's theorem says every point of $\mathcal{C}$ is the barycentre of a probability measure concentrated on $\mathbf{\Lambda}$, and the probability measure is actually unique in our case: for a discussion and proof see [13], Chapter II.

We can summarize the above discussion with the following statement.

Proposition 2.2 (the Godement-Bochner theorem)**.** *Suppose (G, K) is a Gelfand pair, and $\mathbf{\Lambda}$ is the (topological) space of elementary spherical functions for the pair (G, K), or equivalently the (topological) space of equivalence classes of unitary irreducible representations of G having a K-fixed vector. Then the K-bi-invariant states of G are exactly the continuous functions on G which can be written as $\varphi = \int_{\mathbf{\Lambda}} \varphi_\lambda d\mu_\varphi(\lambda)$, where μ_φ is a measure on Λ.*

Let us make the backwards way from the theory of positive-definite functions for a Gelfand pair (G, K) to that of Gaussian random fields on G/K. It starts with a remark: suppose m_1, m_2 are K-bi-invariant states of G and Φ_1, Φ_2 are independent Gaussian fields whose covariance functions, when turned into functions on G as before, are m_1 and m_2, respectively. Then a Gaussian field whose correlation function is $m_1 + m_2$ necessarily has the same probability distribution as $\Phi_1 + \Phi_2$. A simple application of Fubini's theorem extends this remark to provide a *spectral decomposition for Gaussian fields*, which mirrors the above decomposition of spherical functions:

- For every λ in Λ, there is, up to equality of the probability distributions, exactly one Gaussian field whose correlation function is φ_λ ;

- Suppose $(\Phi_\lambda)_{\lambda\in\Lambda}$ is a collection of mutually independent Gaussian fields, and for each λ, Φ_λ has correlation function φ_λ. Then for each probability measure μ on Λ, the covariance function of the Gaussian field
$$x \rightsquigarrow \int_{\mathbf{\Lambda}} \Phi_\Lambda(x) d\mu(\lambda)$$
is $\int_{\mathbf{\Lambda}} \varphi_\lambda d\mu_\varphi(\lambda)$.

In the next section, I will focus on special cases (most importantly, symmetric spaces); in these cases I will give explicit descriptions of Λ and, for each λ in Λ, of the spherical function φ_λ and of a Gaussian field whose correlation function is φ_λ.

3 Existence theorems and explicit constructions

3.1 Semidirect products with a vector normal subgroup: easy no-go results

Suppose H is a Lie group, A is a finite-dimensional vector space, and $\rho : H \to GL(A)$ is a continuous morphism. The semidirect product $G = H \ltimes_\rho A$ (whose underlying set if $H \times A$, and whose composition reads $(h_1, a_1)\cdot(h_2, a_2) := (h_1h_2, a_1+\rho(h_1)a_2)$) is a Lie group.

Since an H-bi-invariant function on G is entirely determined by its restriction to A, the convolution of H-bi-invariant functions on G is a commutative operation. So when H is compact, (G, H) is a Gelfand pair; for this case I shall make the situation fully explicit in subsection 3.3 below. When H is not compact, (G, H) is not a Gelfand pair, and that is not only because the definition as I wrote it needs the compactness: I shall start this section by showing that G-invariant continuous Gaussian fields need not exist on G/H.

Examples. The *Poincaré group* P is the largest subgroup of the affine group of $\mathbb{R}^4$ under which the space of solutions to Maxwell's (that is, the wave) equation for a scalar-valued field in a vacuum is stable. It is a famous result of Poincaré that $P = SO(3,1) \ltimes \mathbb{R}^4$, with the obvious action of $SO(3,1)$ on $\mathbb{R}^4$. The *Galilei* group is the subgroup of the affine group of $\mathbb{R}^4 = \mathbb{R}^3 \times \mathbb{R}$ gathering the transformations which, for pairs of points in $\mathbb{R}^4$, preserve the notion of Euclidean distance between the ("space") projections on $\mathbb{R}^3$ as well as the distance between the ("time") projections on $\mathbb{R}$ (to be formal, the Galilei group consists of affine transformations of $\mathbb{R}^4$ leaving the map $[(x_1,t_1),(x_2,t_2)] \mapsto (\|x_1 - x_2\|, t_1 - t_2)$ from $\mathbb{R}^4 \times \mathbb{R}^4$ to $\mathbb{R}^2$ invariant [10]).

To state our easy no-go result, recall that the Fourier tranform of a function on A is a function on the set $\hat{A}$ of characters of A: . Recall also that an action of H on the Abelian group A yields an action $\hat{\rho}$ of H on $\hat{A}$ if we set $\hat{\rho}(h)\chi := x \mapsto \chi(\rho(h)^{-1}x)$.

Proposition 3.1. *Let G be a semidirect product $H \ltimes A$ as above. If there is no compact orbit of H in $\widehat{A}$ but the trivial one, then no real-valued standard Gaussian field on $\mathbb{R}^4$ whose probability distribution is G-invariant can have continuous trajectories.*

Proof. Such a field would yield a positive-definite, continuous, H-bi-invariant function, say $\tilde{\Gamma}$, on G, taking the value one at 1_G, but we shall see now that there can be no such function except the constant one. Write Γ for the restriction of $\tilde{\Gamma}$ to A; because Γ is positive-definite as a function on the abelian group A, Bochner's theorem (see for instance [2], p. 109) says it is the Fourier transform of a bounded measure ν_Γ on $\widehat{A}$. Because Γ is invariant under the linear action on H on A and because of the elementary properties of the Fourier transform, the measure ν_Γ must also be invariant, and so if Ω is a compact subset of an H-orbit in $\widehat{A}$, $\nu_\Gamma(\Omega)$ must be equal to $\nu_\Gamma(h \cdot \Omega)$ for each h in H. That is not possible when $H \cdot \Omega$ is noncompact unless $\nu_\Gamma(\Omega)$ is zero, because there is a sequence (h_n) in $H^{\mathbb{N}}$ such that $\cup_n h_n \cdot \Omega$ is a disjoint union, and because the total mass of ν_Γ is finite. A consequence is that the support of ν_Γ must be the origin in $\widehat{A}$, and since ν_Γ is the Fourier transform of a continuous function, it must be a multiple of the Dirac mass at zero. $\square$

If $H \ltimes A = SO(3,1) \ltimes \mathbb{R}^4$ is the Poincaré group, we can identify $\widehat{A}$ with $\mathbb{R}^4$ in a H-equivariant way using Minkowski's quadratic form, and then the orbits of H on $\widehat{A}$ appear as the level sets of Minkowski's quadratic form in $\mathbb{R}^4$. So of course the hypothesis of Proposition 3.1 is satisfied:

Corollary. *No standard real-valued Gaussian field on $\mathbb{R}^4$ whose probability distribution is invariant under the Poincaré group can have continuous trajectories.*

Let us now consider whether a standard real-valued Gaussian field on $\mathbb{R}^4$ whose probability distribution is invariant under the Galilei group can have continuous trajectories.

The linear part of an element in the Galilei group reads $(x,t) \mapsto (A\vec{x} + \vec{v}t, t)$, where A is an element of $SO(3)$ and $\vec{v}$ is a vector in $\mathbb{R}^3$, and its inverse reads $(x,t) \mapsto (A^{-1}\vec{x} - (A^{-1}\vec{v})t, t)$; so if χ is in $\widehat{\mathbb{R}^4}$ and decomposes as $(\vec{x},t) \mapsto \langle \vec{k}_\chi, \vec{x}\rangle + \omega_\chi t$, then $\chi(A^{-1}\vec{x} - (A^{-1}\vec{v})t) = \langle A\vec{k}_\chi, \vec{x}\rangle + \left(\omega_\chi + \langle \vec{k}_\chi, \vec{v}\rangle\right) t$. This means that if h is the element $(x,t) \mapsto (A\vec{x} + \vec{v}t, t)$ of

10. The action of the affine group here is the diagonal action on $\mathbb{R}^4 \times \mathbb{R}^4$.

the linear part of the Galilei group,

$$h \cdot \chi := h \cdot (\vec{k}_\chi, v_\chi) = \left(A\vec{k}_\chi, \omega_\chi + \langle \vec{k}_\chi, \vec{v} \rangle \right).$$

The orbits of the Galilei group on $\widehat{\mathbb{R}^4}$ are thus the cylinders $\mathcal{C}_\kappa := \left\{ (\vec{k}, \omega) \mid \left\| \vec{k} \right\| = \kappa \right\}$, $\kappa > 0$, and the points $\{(0, \omega)\}$, $\omega \in \mathbb{R}$. The proof of Proposition 3.1 shows that the support of ν_Γ must be the union of the compact orbits, and this is the "time frequency" axis $\{(0, \omega) \mid \omega \in \mathbb{R}\}$; the measure ν_Γ then appears as the product of the Dirac mass on the line of $\mathbb{R}^4$ which is dual to the "time" axis, with a bounded measure on that line.

So a standard real-valued Gaussian field on $\mathbb{R}^4$ whose probability distribution is invariant under the Galilei group cannot have continuous trajectories without losing any form of space dependence :

Corollary. *A standard, real-valued Gaussian field whose probability distribution is invariant under the Galilei group and which has continuous trajectories reads but $(x, t) \mapsto \Phi(t)$, where Φ is a stationary and continuous Gaussian field on the real line.*

Remark. Proposition 3.1 and its corollaries might seem incompatible with the fact that, leaving Gelfand pairs aside, every unitary representation contributes a continuous positive-definite function. The representation-theoretic counterpart to Proposition 3.1 is thus the fact that no irreducible unitary representation of $H \ltimes A$ except the trivial one can have a H-invariant vector.

If we look for fields with *smooth* trajectories instead of continuous ones [11], an interesting remark by Adler and Taylor makes Proposition 3.1 trivial:

Lemma 3.1. *If there exists a smooth, non-constant, real-valued Gaussian field on G/K whose probability distribution is G-invariant, then K is compact.*

Proof. For each p in G/K and every (X_p, Y_p) in $(T_p(G/K))^2$, set

$$g(X_p, Y_p) := \mathbb{E}\left[(d\Phi(p)X_p)(d\Phi(p)Y_p) \right].$$

This has a meaning as soon as the samples of Φ are almost surely smooth, and it does define a riemannian metric on G/K. The invariance of the field now implies that this metric is G-invariant, and in particular that the positive-definite quadratic form it provides on $T_{x_0}(G/K)$ is K-invariant. Thus K is contained in the isometry group of a finite-dimensional Euclidean space, so it is compact. $\square$

3.2 Monochromatic fields on commutative spaces

Let us start again with a Gelfand pair (G, K) with connected G. From now on, I shall assume that G is a *Lie* group and focus on Gaussian fields which have *smooth* trajectories. The reason, here summarized as Theorem 3.1, is that in this case, the spherical functions are solutions to invariant partial differential equations: as I promised earlier, the coefficients for these partial differential equations must determine all the statistical properties

11. In another direction, one could argue that the inverse Fourier transform of a noncompact orbit, while not a continuous spherical function, is a tempered distribution which could be used to define distribution-valued Gaussian fields on $\mathbb{R}^4$; although we shall not take up this point of view, Yaglom mentioned the possibility at the end of [36], and on the group-theoretic side, the distribution-theoretic theory of *generalized* Gelfand pairs has been worked out, see [11].

the corresponding field, and we shall see this at work with the density of the zero-set. In addition, fully explicit constructions are possible in many cases of interest.

A good reference for this subsection is J. A. Wolf [34].

Let me write $\mathcal{D}_G(X)$ for the algebra of G-invariant differential operators on $X = G/K$. Then Thomas [29] and Helgason [16] proved that (G, K) is a Gelfand pair if and only if $D_G(X)$ is a commutative algebra.

Theorem (See [34], Theorems 8.3.3-8.3.4)**.** *A smooth, K-bi-invariant function ϕ is an elementary spherical function for (G, K) if and only if there is, for each D in $\mathcal{D}_G(X)$, a complex number $\chi(D)$ such that*

$$D\varphi = \chi_\varphi(D)\varphi.$$

The eigenvalue assignation $D \mapsto \chi_\varphi(D)$ defines a character of the commutative algebra $D_G(X)$. It determines the spherical function φ: when χ is a character of $\mathcal{D}_G(X)$, there is a unique spherical function φ such that $\chi = \chi_\varphi$.

Definition. *A standard Gaussian random field on X whose correlation function is a multiple of an elementary spherical function will be called* monochromatic*; the corresponding character of $\mathcal{D}_G(X)$ will be called its* spectral parameter*.*

Note that with a choice of G-invariant riemannian metric on G/K comes an element of $\mathcal{D}_G(X)$, the Laplace-Beltrami operator Δ_X.

Example 3.1. *If X is a* two-point homogeneous space*, that is, if G is transitive on equidistant pairs*[12] *of points in X, then $\mathcal{D}_G(X)$ is the algebra of polynomials in Δ_X.*

Example 3.2. *If X is a* symmetric space *(see below), then $\mathcal{D}_G(X)$ is finitely generated; thus a character of $\mathcal{D}_G(X)$ is specified by a finite collection of real numbers.*

3.3 Explicit constructions. A: Flat homogeneous spaces

Suppose (G, K) is a Gelfand pair, and the commutative space $X = G/K$ is *flat*. Then we know from early work by J. A. Wolf (see [35], section 2.7) that X is isometric to a product $\mathbb{R}^n \times T^s$.

I will be concerned with the simply connected case: let V be a Euclidean space, K be a closed subgroup of $SO(V)$, and G be the semidirect product $K \ltimes V$. Then as we saw (G, K) is a Gelfand pair, and we can describe the G-invariant continuous Gaussian fields on $V = G/K$ from the monochromatic ones.

The next proposition provides a description of the elementary spherical functions.

Proposition 3.2. *Suppose V is a Euclidean vector space, and K is a closed subgroup of $SO(V)$. Then the Fourier transform of a K-orbit in V is a smooth function ; once normalized to take the value one at zero, it is an elementary spherical function for the Gelfand pair $(K \ltimes V, K)$. In fact, every elementary spherical function for $(K \ltimes V, K)$ restricts on V to the Fourier transform of a K-orbit in V.*

12. This means that two pairs of points (p_1, q_1), (p_2, q_2) satisfy $d(p_1, p_2) = d(q_1, q_2)$ if and only if there is an isometry $g \in G$ such that $g \cdot p_1 = q_1$ and $g \cdot p_2 = q_2$.

Proof. To get a handle on the K-bi-invariant states of $K \ltimes V$, let me start with a bounded measure on the orbit space V/K and the measure $\tilde{\mu}$ on V obtained by pulling μ back with the help of the Hausdorff measure of each K-orbit (normalized so that each orbit has total mass one). I first remark that the Fourier transform of $\tilde{\mu}$ provides a positive-definite function for $(K \ltimes V, K)$. Indeed, Bochner's theorem says it provides a positive-definite function on V, and if (k_1, v_1) and (k_2, v_2) are elements of $K \ltimes V$, $(k_1, v_1)(k_2, v_2)^{-1}$ is equal to $(k_1 k_2^{-1}, v_1 - k_1 k_2^{-1} v_2)$, so a K-bi-invariant function takes the same value at $(k_1, v_1)(k_2, v_2)^{-1}$ as it does at $(k_1^{-1}, 0)(k_1, v_1)(k_2, v_2)^{-1}(k_2, 0) = (1_K, k_1^{-1} v_1 - k_2^{-1} v_2)$; this checks the positive-definiteness directly.

Now suppose $\tilde{\Gamma}$ is a K-bi-invariant state of $K \ltimes V$, and write Γ for its restriction to V, a bounded K-invariant positive-definite continuous function on V. The Fourier transform of Γ, a bounded complex measure on V with total mass one because of Bochner's theorem, is K-invariant, and yields a bounded measure on the orbit space K/V. If the support of this measure is not a singleton, we can split it as the half-sum of bounded measures with total mass one, and lifting them to V and taking Fourier transform exhibits our initial state as a sum of two K-bi-invariant functions taking the value one at $1_{K \ltimes V}$ which, according to the previous paragraph, are positive-definite. So the extreme points among the K-bi-invariant states correspond to K-invariant measures concentrated on a single K-orbit in V, which proves the proposition.

$\square$

Remark. Suppose $\mathcal{D}_G(V)$ is finitely generated. Since the elements in $\mathcal{D}_G(V)$ are invariant under the translations of V, they have constant coefficients, so they become multiplication by polynomials after taking Fourier transform. Taking the Fourier transform of what Proposition 3.1 says, we see that a K-orbit in V is an affine algebraic subset of V, and that all orbits are obtained by varying the constant terms in a generating system for the ring of Fourier transforms of elements of $\mathcal{D}_G(V)$. Of course the simplest case is when the orbits are spheres and G is the Euclidean motion group of V.

Proposition 3.2 is explicit enough to allow for computer simulation: suppose φ is an elementary spherical function, and let us see how to build a Gaussian field Φ_Ω on V whose covariance function is φ. By definition, we must have $\mathbb{E}\left[\Phi_\Omega(x)\Phi_\Omega(0)\right] = \varphi(x)$, so using Fubini's theorem we see that (almost) all samples of Φ must have their Fourier transform concentrated on the same K-orbit of V, say Ω, as Φ. Thus Φ is a random superposition of waves whose wave-vectors lie on Ω.

Lemma 3.2. *Assume $\left(\zeta_{\vec{k}}\right)_{\vec{k}\in\Omega}$ is a collection of mutually independent standard Gaussian random variables. Normalize the Hausdorff measure on Ω so that it has total mass one. Then the Gaussian random field*

$$\Phi_\Omega := x \mapsto \left[\int_\Omega e^{i\vec{x}\cdot\vec{k}}\ \zeta_{\vec{k}}\ d\vec{k}\right].$$

is G-homogeneous, smooth, and has covariance function φ_Ω.

Proof. This is straightforward from the definition, since applying Fubini's theorem twice yields

$$
\begin{aligned}
\mathbb{E}\left[\Phi_\Omega(\vec{x})\Phi_\Omega(0)^\star\right] &= \mathbb{E}\left[\left(\int_\Omega e^{i\vec{k}\cdot\vec{x}}\zeta_{\vec{k}}\ d\vec{k}\right)\left(\int_\Omega \zeta^\star_{\vec{u}}\ d\vec{u}\right)\right] \\
&= \mathbb{E}\left[\int_{\Omega^2} e^{i\vec{k}\cdot\vec{x}}\ \zeta_{\vec{k}}\ \zeta^\star_{\vec{u}}\ d\vec{k}d\vec{u}\right] \\
&= \int_{\Omega^2} e^{i\vec{k}\cdot\vec{x}}\ \mathbb{E}\left[\zeta_{\vec{k}}\ \zeta^\star_{\vec{u}}\right]\ d\vec{k}d\vec{u} \\
&= \int_\Omega e^{i\vec{k}\cdot\vec{x}}d\vec{k} \\
&= \varphi_\Omega(\vec{x})
\end{aligned}
$$

as announced. The smoothness and invariance follow from Proposition 2.1. □

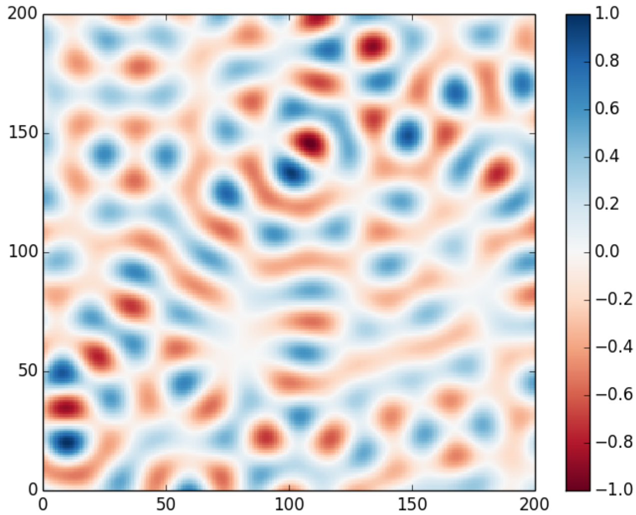


Figure 1: a real-valued map, sampled from a real-valued monochromatic field on the Euclidean plane.

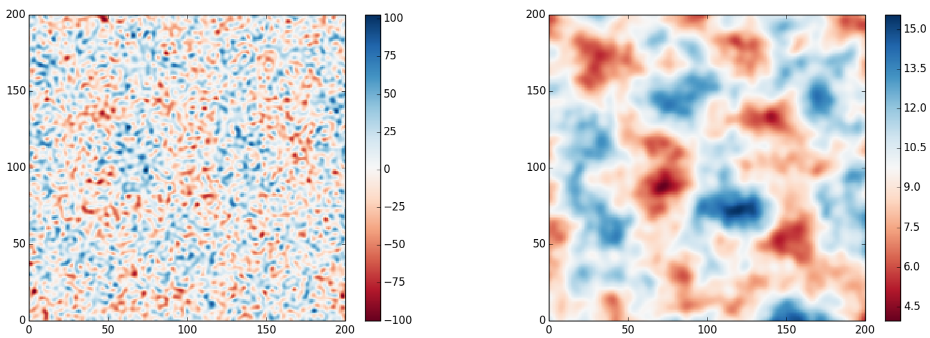


Figure 2: Two real-valued maps, sampled from real-valued invariant fields on the Euclidean plane: because of Proposition 3.2, the elementary spherical functions form a half-line; the power spectrum (the measure on $\mathbb{R}^+$ defined in Proposition 2.2) of the upper map is roughly the indicatrix of a segment, the power spectrum of the lower one has the same support but decreases as $1/R^2$

3.4 Explicit constructions. B: Compact homogeneous spaces

Suppose (G, K) is a Gelfand pair, and the commutative space $X = G/K$ is *positively curved*. Then G is a connected *compact* Lie group, and the Hilbert spaces for irreducible representations of G are finite-dimensional.

A consequence is that if $T : G \to U(\mathcal{H})$ is an irreducible representation, the map $g \mapsto \mathrm{Trace}(T(g))$ is a continuous, complex-valued function; it is of course the *global character* of G.

Proposition 3.3 (G. Van Dijk, see Theorem 6.5.1 in [11]). *The elementary spherical functions for (G, K) are the maps*

$$x \mapsto \int_K \chi(x^{-1}k)dk$$

where χ runs through the set of global characters of irreducible representations of G having a K-fixed vector, and the integration is performed w.r.t the normalized Haar measure of K.

Note that if χ is the global character of an irreducible representation of G which has no K-fixed vector, the above expression is zero.

The reason why this provides an explicit formula for the spherical functions is that Hermann Weyl famously wrote down the global character of an irreducible representation of G. Let T be a maximal torus in G, let $\mathfrak{t}$ and $\mathfrak{g}$ be the complexified Lie algebras of T and G, and let W be the Weyl group of the pair $(\mathfrak{g}, \mathfrak{t})$, $\mathcal{C} \subset \mathfrak{t}^\star$ be a Weyl chamber in $\mathfrak{t}^\star$, Σ be the set of positive roots of $(\mathfrak{g}, \mathfrak{t})$ in the ordering determined by $\mathcal{C}$ − a subset of $\mathfrak{t}^\star$ as well −, ρ be the half-sum of elements of Σ, Λ be the subset of $\mathfrak{t}^\star$ gathering the differentials of continuous morphisms $T \to \mathbb{C}$, and Λ^+ be $\Lambda \cap \mathcal{C}$. Two of the most famous results in representation theory are:

• There is a natural bijection (the highest-weight theory) between Λ^+ and the equivalence classes of irreducible representation of G;

• The global character of all irreducible representations with highest weight λ restricts to $\exp_G(\mathfrak{t})$ as [13]

$$e^H \mapsto \frac{\sum\limits_{w \in W} \varepsilon(w) e^{\langle \lambda+\rho, wH\rangle}}{\sum\limits_{w \in W} \varepsilon(w) e^{\langle \rho, wH\rangle}}.$$

This gives a completely explicit description of the covariance functions of invariant Gaussian random fields on X (provided one can find a maximal torus, the Weyl group, the roots... explicitly: the `atlas` software seems to do that − and much more − when G is reductive). In contrast to what happened above for flat spaces and to what will happen below for symmetric spaces, however, I am not aware that this leads to an explicit description of the Gaussian random field with a given spherical function as its covariance function. We must stick to Yaglom's general construction here: without assuming that (G, K) is a Gelfand pair but only that it is a pair of connected compact Lie groups, let $T : G \to U(\mathcal{H})$ be an irreducible representation, $(\mathbf{e}_1, ...\mathbf{e}_r)$ be an orthonormal basis for the space of K-fixed vectors, and let $(\mathbf{e}_{r+1}, ...\mathbf{e}_d)$ be an orthonormal basis for its orthocomplement. Yaglom proved the following two facts:

13. Recall that the conjugates of $\exp_G(\mathfrak{t})$ is G and the character is conjugation-invariant!

• The maps $gK \mapsto \langle \mathbf{e}_i, T(g)\mathbf{e}_{,}\rangle$, $i, j = 1..r$, are elementary spherical functions for (G, K),
• If $(\zeta_i)_{i=1,\ldots d}$ is a collection of i.i.d. standard Gaussian variables, then for every i_0 in $\{1, ..r\}$,

$$gK \mapsto \sum_{i=1}^{d} \zeta_i \langle \mathbf{e}_i, T(g)\mathbf{e}_{i_0}\rangle \tag{3.1}$$

is an invariant standard Gaussian random field on G/K, whose covariance function is $gK \mapsto \langle \mathbf{e}_1, T(g)\mathbf{e}_1\rangle$.

While Yaglom's result is rather abstract compared to the above descriptions for flat spaces, in many cases of interest explicit bases $(\mathbf{e}_i)$ and explicit formulae for the matrix elements $\langle \mathbf{e}_i, T(g)\mathbf{e}_j\rangle$ are known (the obvious reference is [30]), making (3.1) startingly concrete.

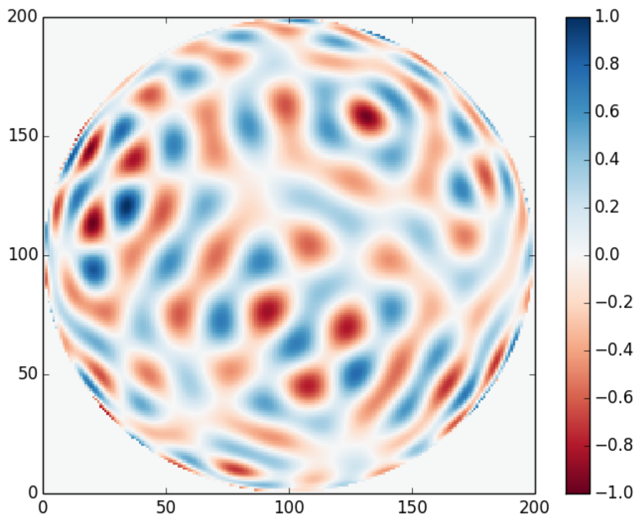


Figure 3: a sample from a real-valued monochromatic field on the sphere. This uses a combination of spherical harmonics with i.i.d. gaussian coefficients.

3.5 Explicit constructions. C: Symmetric spaces of noncompact type

Suppose (G, K) is a Gelfand pair, and the commutative space $X = G/K$ is *negatively curved*. Then G is noncompact, and without any additional hypothesis on G it is quite difficult to do geometry and analysis on X. It is easier to do so if X is a *symmetric* space. The isometry group G is then semisimple.

In that case Harish-Chandra determined the elementary spherical functions for (G, K) in 1958; Helgason later reformulated his discovery in a way which brings it very close to Proposition 3.2. For the contents of this subsection, see chapter III in [17], and see of course [15], [16], [17] for more on the subject.

Suppose G is a connected semisimple Lie group with finite center and K is a maximal compact subgroup in G. Write $\mathfrak{g}$ and $\mathfrak{k}$ for their Lie algebras, $\mathfrak{p}$ for the orthocomplement of $\mathfrak{k}$ with respect to the Killing form of $\mathfrak{g}$, $\mathfrak{a}$ for a maximal abelian subspace of $\mathfrak{p}$. Using a subscript $\cdot_{\mathbb{C}}$ to denote complexifications, let $\mathcal{C} \subset i\mathfrak{a}^\star \subset \mathfrak{a}^\star_{\mathbb{C}}$ be the Weyl chamber corresponding to a choice of positive roots for $(\mathfrak{g}_{\mathbb{C}}, \mathfrak{a}_{\mathbb{C}})$, and ρ be the corresponding half-sum of positive roots. The direct sum of real root spaces for the chosen positive roots is a Lie subalgebra, say $\mathfrak{n}$, of $\mathfrak{g}$, and if A and N are the subgroups $\exp_G(\mathfrak{a})$ and $\exp_G(\mathfrak{n})$ of G, the map $(k, a, n) \mapsto kan$ is a diffeomorphism between $K \times A \times N$ and G. When the Iwasawa decomposition of $x \in G$ accordingly is $k \exp_G(H)n$, let me write $\mathfrak{A}(x) = H$ for the $\mathfrak{a}$-component.

Now suppose λ is in $\mathfrak{a}^\star$ and b is in K. Define

$$\begin{aligned} e_{\lambda,b} : G &\to \mathbb{R} \\ x &\mapsto e^{\langle i\lambda+\rho \ \mid \ \mathfrak{A}(b^{-1}\tilde{x})\rangle}. \end{aligned}$$

Then $e_{\lambda,b}$ defines a smooth function from $X = G/K$ to $\mathbb{C}$; it is an eigenfunction of Δ_X, with eigenvalue[14] $-\left(\||\lambda\|^2 + \|\rho\|^2\right)$. These functions are useful for harmonic analysis on G/K in about the same way as plane waves are useful for classical Fourier analysis.

If (λ_1, b_1) and (λ_2, b_2) are elements of $\mathfrak{a}^\star \times K$, then e_{λ_1,b_1} and (λ_2, b_2) coincide if and only if there is an element w in the Weyl group of $(\mathfrak{g}_{\mathbb{C}}, \mathfrak{a}_{\mathbb{C}})$ such that $\lambda_1 = w\lambda_2$ and if b_1 and b_2 have the same image in the quotient $B = K/M$, where M is the centralizer of $\mathfrak{a}$ in K. Each of the $e_{\lambda,b}$ thus coincides with exactly one of the $e_{\lambda^+,b}$s, where λ^+ runs through the closure Λ^+ of $\mathcal{C}$ in $i\mathfrak{a}^\star$.

Theorem (Harish-Chandra)**.** *For each λ in Λ^+,*

$$\varphi_\lambda := x \mapsto \int_B e_{\lambda,b}(x) db$$

is[15] an elementary spherical function for (G, K). Every spherical function for (G, K) is one of the φ_λ, $\lambda \in \Lambda^+$.

Thus the possible spectral parameters for monochromatic fields occupy a closed cone Λ^+ in the Euclidean space $i\mathfrak{a}^\star$ (and the topology on the space of spherical functions described in section 2.2 coincides with the topology inherited from $\mathfrak{a}^\star$). The fact that spherical functions here again appear as a constructive interference of waves yields an explicit description for the monochromatic field with spectral parameter λ (same proof as Lemma 3.2):

Lemma 3.3. *Assume $(\zeta_b)_{b\in B}$ is a collection of mutually independent standard Gaussian random variables. Then the Gaussian field*

$$\Phi_\lambda := x \mapsto \left[\int_B e_{\lambda,b}\ \zeta_b\ db\right].$$

is G-homogeneous, smooth, and has covariance function φ_λ.

14. Here the norm is the one induced by the Killing form.
15. Here the invariant measure on B is normalized so as to have total mass one.

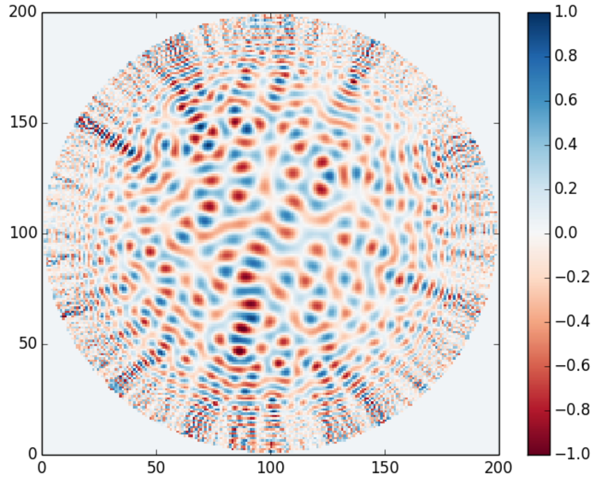


Figure 4: a real-valued map on the Poincaré disk, sampled from a monochromatic field using Lemma 3.3.

4 The typical spacing in an invariant field

Let us start with a homogeneous *real*-valued Gaussian field Φ on a riemannian homogeneous space X with isometry group G. In view of the above pictures, if the correlation function of Φ is close enough to being an elementary spherical function, one expects Φ to exhibit some form of quasiperiodicity [16].

Let us now see whether we can give a meaning to the "quasiperiod". Draw a geodesic γ on X, and if Σ is a segment on γ, write $\mathcal{N}_\Sigma$ for the random variable recording the number of zeroes of Φ on Σ. Because the field Φ is homogeneous and the metric on X is invariant, the probability distribution of $\mathcal{N}_\Sigma$ depends only on the length, say $|\Sigma(\gamma)|$, of Σ. The identity component of the subgroup of G fixing γ is a one-parameter subgroup of G, and reads $\exp_G(\mathbb{R}\vec{\gamma})$ for some $\vec{\gamma}$ in $\mathfrak{g}$; it is isomorphic to a circle if X is of the compact type, and isomorphic to the additive group of the real line if X is of the Euclidan or noncompact type. In any case, this means we can pull back $\Phi|_\gamma$ to $\mathbb{R}\vec{\gamma}$ and view it as a stationary, real-valued Gaussian field on the real line. In this way, the group exponential relating $\mathbb{R}\vec{\gamma}$ to γ sends the Lebesgue measure of $\mathbb{R}$ to a constant multiple of the metric γ inherits from that of X. The zeroes of the pullback of $\Phi|_\gamma$ to $\mathbb{R}\vec{\gamma}$ can thus be studied through the classical, one-dimensional, Kac-Rice formula:

Proposition (Rice's formula)**.** *Suppose Φ is a translation-invariant smooth Gaussian field on the real line, with smooth trajectories ; choose a real number u, and consider an interval I of length ℓ on the real line. Write $\mathcal{N}_{u,I}$ for the random variable recording the number of points x on I where $\Phi(x) = u$; then*

$$\mathbb{E}\left[\mathcal{N}_{u,I}\right] = \ell \cdot \frac{e^{-u^2/2}\sqrt{\lambda}}{\pi} \tag{4.1}$$

where $\lambda = \mathbb{E}\left[\Phi'(0)^2\right]$ is the second spectral moment of the field.

16. In a mathematically loose sense.

An immediate consequence of (4.1) is that the expectation $\mathbb{E}[\mathcal{N}_\Sigma]$ depends linearly on $|\Sigma|$.

Definition. *The* typical spacing *of Φ is the positive number $\Lambda(\Phi)$ such that*

$$\frac{1}{\Lambda(\Phi)} := \frac{\mathbb{E}[\mathcal{N}_\Sigma]}{|\Sigma(\gamma)|}.$$

For a comment on the definition, see Example 4.2 below.

Proposition 4.1. *Suppose X is a riemannian homogeneous space, and in the setting of Definition 5.1, assume the samples of Φ lie almost surely in the eigenspace $\{f \in \mathcal{C}^\infty(X) \mid \Delta_X f = Kf\}$, and write β for the variance of $\Phi(x)$ at any point $x \in X$. Then*

$$\Lambda(\Phi) = \frac{\pi}{\sqrt{(dim\ X)\beta|K|}}.$$

Proof. Let me write κ for the second spectral moment of the stationary gaussian field on the real line, say $\mathbf{u}$, obtained by restricting Φ_γ to $\mathbb{R}_{\vec{\gamma}}$ as above: κ is the variance $\mathbb{E}\left[\mathbf{u}'(0)^2\right]$. Because of (4.1), $\Lambda(\Phi)$ is equal to $\frac{\pi}{\sqrt{\kappa}}$.

Now, $\mathbf{u}'(0)$ is the derivative of Φ in the direction $\vec{\gamma}$. Its variance can be recovered from the second derivative of the covariance function of Φ in the direction $\vec{\gamma}$: let me write Γ for the covariance function of Φ, turned into a function on G thanks to a choice of base point x_0 in X. Recall that $\Gamma(a^{-1}b) = \mathbb{E}[\Phi(a\cdot x_0)\Phi(b\cdot x_0)]$, consider the functions $f_1 : (a,b) \to \Gamma(a^{-1}b)$ and $f_2 : (a,b) \to \mathbb{E}[\Phi(a\cdot x_0)\Phi(b\cdot x_0)]$ from G^2 to $\mathbb{C}$. Write the Lie derivative in the direction $\vec{\gamma}$ with respect to a or b as as $L^a_{\vec{\gamma}}$ or $L^b_{\vec{\gamma}}$. Then $\left(L^a_{\vec{\gamma}}L^b_{\vec{\gamma}}f_1\right)(1_G) = -L^2_{\vec{\gamma}}(\Gamma)(1_G)$, while $\left(L^a_{\vec{\gamma}}L^b_{\vec{\gamma}}f_2\right)(1_G) = \mathbb{E}\left[\left(L_{\vec{\gamma}}\Phi\right)(x_0)^2\right]$. Naturally $f_1 = f_2$, so

$$\mathbb{E}\left[\left(L_{\vec{\gamma}}\Phi\right)(x_0)^2\right] = -L^2_{\vec{\gamma}}(\Gamma)(1_G).$$

If Γ_X is the map $x \mapsto \Gamma(x)$, then $-L^2_{\vec{\gamma}}(\Gamma_X)(x_0)$. Of course, the Laplace-Beltrami operator on X has much to do with second derivatives :

• when X is flat, Δ_X is the usual laplacian, we can choose Euclidean coordinates on X such that $\mathbb{R}\vec{\gamma}$ is the first coordinate axis; writing X_i for the vector fields generating the translations along the coordinate axes, we then have $\Delta_X = \sum_{i=1}^{(\text{dim X})} L^2_{X_i}$.

• In the general case, we can localize the computation and use normal coordinates around x_0: suppose $(\gamma_1^{x_0}, ...\gamma_p^{x_0})$ is an orthonormal basis of $T_{x_0}X$, and let $\vec{\gamma}_1, ...\vec{\gamma}_p$ be elements of $\mathfrak{g}$ whose induced vector fields on X coincide at x_0 with the γ_is. Then $(\Delta_X\Gamma)(x_0) = \sum_{i=1}^{p}\left(L_{\vec{\gamma}_i}\Gamma\right)(x_0)$.

We now use the fact that the field is G-invariant and note that the directional derivatives of Γ_X at the identity coset are all identical ; so

$$\left(L^2_{\vec{\gamma}}\Gamma\right)(x_0) = \frac{1}{\dim X}\left(\Delta_X\Gamma_X\right)(x_0).$$

In the special case where Γ is an eigenfunction of Δ_X, we thus get

$$\kappa = (\text{dim X})|K|\Gamma(0) = (\text{dim X})\beta|K|$$

(recall that K is nonpositive when X is compact and nonnegative otherwise), and Proposition 4.1 follows. $\square$

Example 4.1. *Suppose X is the Euclidean plane, and we start from the monochromatic complex-valued invariant field, say $\mathbf{\Phi}$, with characteristic wavelength λ. Then its real part $\Phi_{\mathbb{R}}$ has $\beta = 1/2$ and $\Lambda(\Phi_{\mathbb{R}}) = \lambda$. This we may have expected, since the samples of $\mathbf{\Phi}$ are superpositions of waves with wavelength λ.*

When the curvature is nonzero, however, Proposition 5.1 seems to say something nontrivial.

Example 4.2. *Suppose X is a symmetric space of noncompact type, and we start from a monochromatic invariant field, say Φ, with spectral parameter ω and $\beta = 1/(\dim X)$. In the notations of section 3.4, we get*

$$\Lambda(\mathbf{\Phi}) = \frac{2\pi}{\sqrt{|\omega|^2 + |\rho|^2}}.$$

This is not quite as unsurprising as Example 4.1 : the samples of $\mathbf{\Phi}$ are superpositions of Helgason waves whose phase surfaces line up at invariant distance $\frac{2\pi}{|\omega|}$. The curvature-induced shift in the typical spacing comes from to the curvature-induced growth factor in the eigenfunctions for Δ_X.

Example 4.3. *Suppose X is a compact homogeneous space. Then the gap between zero and the first nonzero eigenvalue*[17] *of Δ_X provides a nontrivial upper bound for the typical spacing of invariant gaussian fields on X (this upper bound is* not *the diameter of X).*

This is clear from Lemma 4.1 for fields with samples in an eigenspace of Δ_X, and the next lemma will make it clear for other fields also.

For general invariant fields on commutative spaces, we can recover the typical spacing as follows:

Lemma 4.1. *Suppose X is a commutative space, Φ is a smooth, invariant, real-valued Gaussian field on X, and write β for the variance of $\Phi(x)$ at any point $x \in X$. Write the spectral decomposition of Φ (section 2) as*

$$\Phi = \int_\Lambda \Phi_\lambda dP(\lambda);$$

then

$$\left(\frac{2\pi}{\Lambda(\Phi)}\right)^2 = \int_\Lambda \left(\frac{2\pi}{\Lambda(\Phi_\lambda)}\right)^2 dP(\lambda).$$

Proof. Let me write Γ for the covariance function of Φ, φ_λ for the spherical function with spectral parameter λ. Note that $\Gamma = \int_\Lambda \varphi_\lambda dP(\lambda)$ as we saw, and taking up the notations of the proof of Lemma 4.1, recall that

$$\left(\frac{2\pi}{\Lambda(\Phi)}\right) = L^2_{\vec{\gamma}}\left(\Gamma\right).$$

I just need to evaluate $L^2_{\vec{\gamma}}\left(\Gamma\right)$. But of course switching with the integration with respect to λ yields

$$L^2_{\vec{\gamma}}\left(\Gamma\right)(x_0) = \int_\Lambda L^2_{\vec{\gamma}}\left(\varphi_\lambda\right)(x_0) dP(\lambda),$$

and the lemma follows. □

17. Relating this to the geometry of X is a deep question ! See for instance [9], III.D.

Remark. The hypotheses in Lemma 5.2 are of course unnecessarily stringent given the proof, and one can presumably evaluate the typical spacing of a general field on a riemannian homogeneous space X by using spectral theory to split it into fields with samples in an eigenspace of Δ_X.

5 Density of zeroes for invariant smooth fields on homogeneous spaces

5.1 Statement of the result

In this section, the homogeneous space X need not be commutative, but need only be riemannian.

Let us start with a definition. Suppose $\mathbf{\Phi}$ is an invariant Gaussian field on X with values in a finite-dimensional vector space V. For each u in V, the typical spacing $\Lambda\left(\langle u|\mathbf{\Phi}\rangle\right)$ of the projection of Φ on the axis $\mathbb{R}u$ depends on the variance β_u of the real-valued Gaussian variable $\Lambda\left(\langle u|\mathbf{\Phi}(p)\rangle\right)$ (here p is any point of X), but $\sqrt{\beta_u}\Lambda\left(\langle u|\mathbf{\Phi}\rangle\right)$ does not depend on u. Choosing an orthonormal basis $(u_1, ...u_{\dim V})$ of V, we can form the quantity $\prod_{i=1}^{\dim V} \sqrt{\beta_u}\Lambda\left(\langle u|\mathbf{\Phi}\rangle\right)$; it does not depend on the chosen basis, I will call it the *volume of an elementary cell for* Φ, and write $\mathcal{V}(\Phi)$ for it.

The terminology is transparent if $\dim V$ and $\dim X$ coincide, provided $\Phi(p)$ is an isotropic Gaussian vector and β_u equals 1 for each u. The notion corresponds to the notion of *hypercolumn* from neuroscience (see [19] for the biological definition, [32] for its geometrical counterpart).

Theorem 5.1. *Suppose $\mathbf{\Phi}$ is a smooth, invariant Gaussian random field on X with values in $\mathbb{R}^{\dim X}$. Write $\mathcal{N}_A$ for the random variable recording the number of zeroes of $\mathbf{\Phi}$ in a Borel region A of X, and Vol(A) for its volume (measured using the G-invariant metric introduced above). Write $\mathcal{V}(\mathbf{\Phi})$ for the volume of an elementary cell for Φ. Then*

$$\mathbb{E}\left(\mathcal{N}_A\right)\frac{\mathcal{V}(\mathbf{\Phi})}{\mathit{Vol}(A)} = (\dim X)!\left(\frac{\pi}{2}\right)^{(\dim X)/2}.$$

Remark. My reason for stating Theorem 5.1 on its own, even though it is a special case of Theorem 3 below, is that the two-dimensional result which motivated this study is one in which it is natural to have $\dim X = 2$ and $V = \mathbb{C}$, and that Theorem 5.1 is a neatly stated generalization to higher dimensions.

Remark. When X is a Euclidean space, the set $\widehat{X}$ of plane waves naturally identifies with X, and so a map from X to $\widehat{X}$ can be interpreted as an arrangment, on X, of labels for Fourier coefficients − very convenient when one remembers that each visual cortical neuron specializes in a single Fourier coefficient of the visual input, and that the various neurons have to be arranged on the (sometimes approximately plane) cortical surface.

When $X = G/K$ is a riemannian symmetric space of the noncompact type, it is true also that a map from X to a Euclidean space with the same dimension can be interpreted as an arrangment of waves on X. Let me use the notations of section 3.3. The standard structure theory of semisimple Lie groups (see e.g. [22], chapter V) says that the "polar coordinates" map $(b, \lambda) \mapsto Ad^\star(b) \times \lambda$ induces a bijection between $B \times (\Lambda/W)$ and $\mathfrak{p}^\star - \{0\}$. This makes it possible to interpret a map from X to $\mathfrak{p}^\star$ − whose dimension is that of X −

as an *arrangment of Helgason waves* on X, with the zeroes corresponding to singularities where the "polar coordinates" are not defined.

⋆

Theorem 5.1 can be extended to a result on the volume of the zero-set of Gaussian fields with values in a Euclidean space of any dimension, as follows. If $\mathbf{\Phi}$ is a smooth invariant Gaussian field on a symmetric space X with values in a finite-dimensional space V, then the zero-set of Φ is generically a union of $(\dim X - \dim V)$-dimensional submanifolds (and is generically empty if $\dim V > \dim X$. Every submanifold of X inherits a metric, and hence a volume form, from that of X, and this almost surely gives a meaning to the volume of the intersection of $\Phi^{-1}(0)$ with a compact subset of X. When A is a Borel region of X and u is an element of V, we wan thus define a real-valued random variable $\mathcal{M}_{\mathbf{\Phi},A}(u)$ by recording the volume of $A \cap \mathbf{\Phi}^{-1}(u)$ for all samples of $\mathbf{\Phi}$ for which u is a regular value, and recording, say, zero for all samples of $\mathbf{\Phi}$ for which u is a singular value.

Theorem 5.2. *Suppose $\mathbf{\Phi}$ is a reduced invariant Gaussian random field on a homogeneous space X with values in a Euclidean space V. Write $\mathcal{M}_{\mathbf{\Phi},A}$ for the random variable recording the geometric measure of $\mathbf{\Phi}^{-1}(0)$ in a Borel region A of X, and Vol(A) for the volume of A. Write $\mathcal{V}(\mathbf{\Phi})$ for the volume of an elementary cell for Φ. Then*

$$\mathbb{E}\left(\mathcal{M}_{\mathbf{\Phi},A}\right) \cdot \frac{\mathcal{V}(\mathbf{\Phi})}{Vol(A)} = \frac{(\dim X)!}{(\dim X - \dim V)!} \cdot \left(\frac{\pi}{2}\right)^{(\dim V)/2}$$

Theorem 2 obviously implies Theorem 1 if we take as a convention that $\mathcal{M}_{\mathbf{\Phi},A}(u)$ is $N(A,u)$ when $\dim X$ and $\dim V$ coincide.

Remark. Thus, in the unit provided by the volume of an elementary cell, the density of the zero-set in an invariant field depends only on the dimension of the source and target spaces, and *not on the group acting.* Of course the group structure is quite relevant for determining the appropriate unit, as we saw.

Remark. I should remark here that when $\dim X$ and $\dim V$ do not coincide, the volume unit $\mathcal{V}(\mathbf{\Phi})$ is not the volume of anything $\dim X$-dimensional in any obvious way – but $\mathbb{E}\left(\mathcal{M}_{\mathbf{\Phi},A}\right)$ is not, either. It is Theorem 5.2 that makes it natural to interpret $\mathcal{V}(\mathbf{\Phi})$ as a volume unit.

5.2 Proof of Theorem 5.2

I will use Azais and Wschebor's Kac-Rice formula for random fields (Theorem 6.8 in [5]); the proof of Theorem 5.2 will be a rather direct adaptation of the one which appears for complex-valued fields on the Euclidean plane and space in [6], [7].

Let me recall their formula, adding a trivial adaptation to our situation where the base space is a riemannian manifold rather than a Euclidean space.

Lemma 5.1. *Suppose (M,g) is a riemannian manifold, and $\Phi : M \rightsquigarrow \mathbb{R}^{\dim M}$ is a smooth Gaussian random field. Assume that the variance of the Gaussian vector $\Phi(p)$ at each point p in M is nonzero.*
For each u in $\mathbb{R}^{\dim M}$ and every Borel subset A in M, write $N(A,u)$ for the random variable recording the number of points in $\Phi^{-1}(u)$.
Then as soon as $\mathbb{P}\left\{\exists p \in M \ , \ \Phi(p) = u \text{ and } \det\left[d\Phi(p)\right] = 0\right\} = 0$,

$$\mathbb{E}\left[N(A,u)\right] = \int_A \mathbb{E}\left\{|\det\left[d\Phi(p)d\Phi(p)^\dagger\right]|^{1/2}\ ||\Phi(p)=u\right\} p_{\Phi(p)}(u)dVol_g(p) \tag{5.1}$$

Proof. After splitting A into a suitable number of Borel subsets, I can obviously work in a single chart and assume that A is contained in an open subset U of M for which there is a diffeomorphism $\psi : M \supset U \to \psi(U) \subset \mathbb{R}^{\dim M}$. I turn $\Phi|_U$ into a Gaussian random field Ψ on $\mathbb{R}^{\dim M}$ by setting

$$\Psi \circ \psi = \Phi.$$

Then I can apply Theorem 6.2 in [5] to count the zeroes of Ψ in $\psi(A)$; since there are as many zeroes of Ψ in $\psi(A)$ as there are zeroes of Φ in A, the theorem yields

$$\mathbb{E}\left[N(A,u)\right] = \int_{\psi(A)} \mathbb{E}\left\{|\det\left[d\Psi(x)d\Psi(x)^\dagger\right]|^{1/2}\ ||\Psi(x)=u\right\} p_{\Psi(x)}(u)dx,$$

where the volume element is Lebesgue measure.

Now, let us start from the right-hand-side of (5.1) and change variables using ψ; we get

$$\int_A \mathbb{E}\left\{|\det\left[d\Phi(p)d\Phi(p)^\dagger\right]|\ |\Phi(p)=u\right\} p_{\Phi(p)}(u)dVol_g(p) =$$ [18]

$$\int_{\psi(A)} \mathbb{E}\left\{|\det\left[d\Phi(\psi^{-1}(x))d\Phi(\psi^{-1}(x))^\dagger\right]|^{1/2}\ |\Phi(\psi^{-1}(x))=u\right\} p_{\Phi(\psi^{-1}(x))}(u)\left|\det\left[d\psi^{-1}(x)\right]\right| dx =$$ [19]

$$\int_{\psi(A)} \mathbb{E}\left\{|\det\left[d\psi^{-1}(x)\right]\det\left[d\Phi(\psi^{-1}x)\right]\det\left[d\Phi(\psi^{-1}x)^\dagger\right]|\det\left[d\psi^{-1}(x)^\dagger\right]|^{1/2}\ |\Psi(x)=u\right\} p_{\Psi(x)}(u)dx =$$ [20]

$$\int_{\psi(A)} \mathbb{E}\left\{|\det\left[d\Psi(x)d\Psi(x)^\dagger\right]|^{1/2}\ |\Psi(x)=u\right\} p_{\Psi(x)}(u)dx = N(A,u)$$

as announced. □

Let us return to the case where Φ is an invariant Gaussian field on a homogeneous space. Choose an orthonormal basis $(u_1,...u_{\dim V})$ of V, write β_i for the standard deviation of the Gaussian variable $\langle u_i, \mathbf{\Phi}(p)\rangle$ at each p (which does not depend on p, and $\mathcal{V}$ for the quantity $\beta_1...\beta_{\dim V}$, which is the volume of the characteristic ellipsoid for the Gaussian vector $\Phi(p)$ at each p and depends neither on p nor on the choice of basis in V.

To prove Theorem 2 we need to look for for $N(A,0)$, and since the field $\mathbf{\Phi}$ is Gaussian, we know that $p_{\Phi(p)}(0) = \mathcal{V}(2\pi)^{-(\dim V)/2}$ for each p. In addition, because of the invariance we know that $p \mapsto \mathbb{E}\left[\Phi(p)^2\right]$ is a constant function on X, so for any vector field $\vec{\gamma}$ on X,

$$\mathbb{E}\left[(L_{\vec{\gamma}}\Phi)(p)\Phi(p)\right] = 0.$$

A first consequence is that $\mathbb{P}\left\{\exists p \in M\ ,\ \Phi(p) = 0 \text{ and } \det\left[d\Phi(p)d\Phi(p)^\dagger\right] = 0\right\}$ is indeed zero, and that we can use Lemma 5.1. Another consequence is that if we choose a basis in T_pX and view $d\Phi(p)$ as a matrix, the entries will be Gaussian random variables which are independent from every component of $\Phi(p)$. This means we can remove the conditioning in (5.1). Thus,

$$\mathbb{E}\left[N(A,0)\right] = \frac{1}{(2\pi\mathcal{V}^2)^{(\dim X)/2}} \int_A \mathbb{E}\left\{|\det\left[d\Phi(p)d\Phi(p)^\dagger\right]|\right\} dVol_g(p). \tag{5.2}$$

Now, $d\Phi(p)$ is a random endomorphism from T_pX to V. Recall that if γ is a tangent vector to X at p, the probability distribution of $(L_\gamma\Phi)(p)$, a Gaussian random vector in V, does not depend on p, and does not depend on γ. Thus there is a basis $(v_1, ...v_{\dim V})$ of V such that for each γ in T_pX, $\langle (L_\gamma\Phi)(p), v_i\rangle$ is independent from $\langle (L_\gamma\Phi)(p), v_j\rangle$ if $i \neq j$ (the v_is generate the principal axes for $(L_\gamma\Phi)(p)$). If we choose any basis of T_pX and write down the corresponding matrix for $d\Phi(p)$ (it has $\dim X$ rows and $\dim V$ columns), then the columns will be independent and will be isotropic Gaussian vectors in $\mathbb{R}^{\dim X}$.

To go furher, we need the following simple remark.

Lemma 5.2. *Suppose M is a matrix with n rows and k columns, $n \geq k$, and write $(m_1, ...m_k) \in (\mathbb{R}^n)^k$ for its columns. Then the determinant of $MM^\dagger$ is the square of the volume of the parallelotope* $\left\{\sum_{i=1}^k t_i m_i \mid t_i \in [0,1]\right\}$.

Proof. Choose an orthonormal basis $(m_{k+1}, ...m_n)$ of $\mathrm{Span}(m_1, ...m_k)^\perp$. Then the signed volume of the k-dimensional parallelotope $\left\{\sum_{i=1}^k t_i m_i \mid t_i \in [0,1]\right\}$ is the same as that of the n-dimensional parallelotope $\left\{\sum_{i=1}^n t_i m_i \mid t_i \in [0,1]\right\}$.

Write $\tilde{M}$ for the $n \times n$ matrix whose columns are the coordinates of the m_i in the canonical basis of $\mathbb{R}^n$. Then $\tilde{M}\tilde{M}^\dagger$ is block-diagonal, one block is $MM^\dagger$ and the other block is the identity because $(m_{k+1}, ...m_n)$ is an orthonormal family.

Thus the determinant of $MM^\dagger$ is the square of that of $\tilde{M}$, and $\det(\tilde{M})$ is the volume of the parallelotope $\left\{\sum_{i=1}^n \alpha_i m_i \mid \alpha_i \in [0,1]\right\}$. □

Coming back to the proof of Theorem 2, we are left with evaluating the mean Hausdorff volume of the random parallelotope generated by $\dim V$ independent isotropic Gaussian vectors in $\mathbb{R}^{\dim X}$.

Lemma 5.3. *Suppose $u_1, ...u_k$ are independent isotropic Gaussian vectors with values in $\mathbb{R}^n$, so that the probability distribution of u_i is $x \mapsto \frac{1}{\alpha_i\sqrt{2\pi}} e^{-\|x\|^2/2\alpha_i^2}$. Write $\mathcal{V}$ for the characteristic volume $\alpha_1...\alpha_k$, and write $\mathbf{V}$ for the random variable recording the k-dimensional volume of the parallelotope $\left\{\sum_{i=1}^k t_i u_i \mid t_i \in [0,1]\right\}$. Then*

$$\mathbb{E}[\mathbf{V}] = \frac{n!}{(n-k)!}\mathcal{V}.$$

Proof. Let me start with k (deterministic) vectors in $\mathbb{R}^n$, say $u_1^0, ...u_k^0$, and choose a basis $u_{k+1}^0, ...u_n^0$ for $\mathrm{Span}(u_1^0, ...u_n^0)^\perp$ Since $\det(u_1^0, ...u_k 0) = \det(u_1^0, ...u_n 0)$ is the (signed) volume of the parallelotope generated by the u_i^0s, we can use the "base times height" formula: writing P_V for the orthogonal projection from $\mathbb{R}^n$ onto a subspace V,

$$\mathrm{Vol}(u_1^0, ...u_n^0) = \left\| P_{\mathrm{Span}(u_2^0, ...u_n^0)^\perp}(u_1^0)\right\| \mathrm{Vol}\left(u_2^0, ...u_n^0\right).$$

Of course then

$$\mathrm{Vol}(u_1^0, ...u_n^0) = \prod_{i=1}^k \left\| P_{\mathrm{Span}(u_{i+1}^0, ...u_n^0)^\perp}(u_i^0)\right\|.$$

Let me now return to the situation with random vectors. Because $u_1, ...u_k$ are independent, the above formula becomes

$$\mathbb{E}\left[\mathrm{Vol}(u_1, ...u_k)\right] = \prod_{i=1}^{k} \mathbb{E}\left[N(u_i, V^i)\right]$$

where $N(u_i, V^i)$ is the random variable recording the norm of the projection of u_i on any (i)-dimensinal subspace of $\mathbb{R}^n$. The projection is a Gaussian vector, and so its norm has a chi-squared distribution with i degrees of freedom. Given the probability distribution of u_i, the expectation for the norm is then $i\alpha_i$, and this does prove Lemma 5.3. □

To complete the proof of Theorem 5.2, choose an orthonormal basis $(\gamma_1, ...\gamma_n)$ in T_pX. Apply Lemma 5.3 to the family $\left(\left(L_{\gamma_i}\langle v_i, \mathbf{\Phi}\rangle\right)(p)\right)_{i=1..n}$. Then (5.2) becomes

$$\mathbb{E}\left[N(A,0)\right] = \frac{1}{(2\pi)^{(\dim V)/2}\mathcal{V}}\mathrm{Vol}(A)\frac{(\dim X)!}{(\dim X - \dim V)!}\prod_{i=1}^{\dim V}\mathbb{E}\left[(L_{\gamma_1}(\langle v_i, \mathbf{\Phi}\rangle)(x_0))^2\right]^{1/2}.$$

To bring the typical spacing back into the picture, recall that the definition and the Kac-Rice formula (4.1) say that $\mathbb{E}\left[(L_{\gamma_1}(\langle v_i, \mathbf{\Phi}\rangle)(x_0))^2\right]^{1/2}$ is none other than $\frac{\pi}{\Lambda(\langle v_i, \mathbf{\Phi}\rangle)}$. Thus

$$\mathbb{E}\left[N(A,0)\right]\frac{\mathcal{V}\prod_{i=1}^{d}\Lambda(\langle v_i, \mathbf{\Phi}\rangle)}{\mathrm{Vol}(A)} = \frac{\pi^{\dim V}}{(2\pi)^{(\dim V)/2}}\frac{(\dim X)!}{(\dim X - \dim V)!}.$$

and since $\mathcal{V}\prod_{i=1}^{d}\Lambda(\langle v_i, \mathbf{\Phi}\rangle)$ is the volume of an elementary cell for $\mathbf{\Phi}$, Theorem 2 is established.

□

Bibliography

[1] P. Abrahamsen, *A review of Gaussian random fields and correlation functions.* Norsk Regnesentral, 1995.

[2] *R. J. Adler* & *J. E. Taylor*, *Random fields and geometry* Berlin: Springer; 2009.

[3] A. Afgoustidis, *Orientation Maps in V1 and non Euclidean Geometry*, The Journal of Mathematical Neuroscience **5**:12 (2015)

[4] A. Afgoustidis, *Monochromaticity of Orientation Maps in V1 Implies Minimum Variance for Hypercolumn Size*, The Journal of Mathematical Neuroscience **5**:10 (2015)

[5] J. M. Azais & M. Wschebor, *Level Sets and Extrema of Random Processes and Fields*, Wiley, 2009.

[6] J. M. Azais, J.R. Leon & M. Wschebor, *Rice formulae and Gaussian waves.* Bernoulli **17(1)** (2011), 170-193.

[7] J. M. Azais, J.R. Leon & M. Wschebor, *Rice formulae and Gaussian waves II.* Publicationes Matematicas del Uruguay **12** (2011), 15-38.

[8] P. Baldi, D. Marinucci & V. S. Varadarajan, *On the characterization of isotropic Gaussian fields on homogeneous spaces of compact groups*, Electron. Comm. Probab. **12** (2007), 291-302.

[9] M. Berger, P. Gauduchon & P. Mazet, *Le spectre d'une variété riemanienne*, Springer Lecture Notes in Mathematics, Vol. 194 (1971)

[10] M. V. Berry & M. R. Dennis, *Phase singularities in isotropic random waves*, Proc. Roy. Soc. London A **456** (2001), 2059-2079.

[11] G. Van Dijk, *Introduction to Harmonic Analysis and Generalized Gelfand Pairs*, de Gruyter Studies in Mathematics, vol. 39 (2009)

[12] S. Cohen & M. A. Lifshits, *Stationary Gaussian Random Fields on Hyperbolic Spaces and on Euclidean Spheres*, ESAIM: Probability and Statistics / Volume 16 / 2012, pp 165 – 221

[13] J. Faraut, *Finite and infinite dimensional spherical analysis.* Summer School: Large N limits, l'étang de Hasselfurth (2008)

[14] B. Galerne, Y. Gousseau & J. M. Morel , *Random Phase Textures: Theory and Synthesis*, IEEE Transactions in Image Processing **20(1)** (2011), 257-267.

[15] S. Helgason, *Differential geometry, Lie groups, and symmetric spaces.* New York: Academic press; 1979.

[16] S. Helgason, *Groups & Geometric Analysis: Radon Transforms, Invariant Differential Operators and Spherical Functions* New York: Academic Press; 1984

[17] S. Helgason, *Geometric analysis on symmetric spaces.* American Mathematical Society; 2008.

[18] T. Hida & T. Hitsuda, *Gaussian Processes*, AMS translations of mathematical monographs, 1990.

[19] D. Hubel, *Eye, brain, and vision.* New York : Scientific American Library, 1995.

[20] A. Malyarenko, *Invariant Random Fields on Spaces with a Group Action*, Probability and Its Applications, Springer (2013)

[21] D. Marinucci & G. Peccati, *Random Fields on the Sphere: Representations, Limit Theorems and Cosmological Applications*, Cambridge University Press (2011)

[22] A. W. Knapp, *Representation theory of semisimple groups: An overview based on examples*, 2nd ed., Princeton university press, Princeton 2001.

[23] J. Mather, *From the Big Bang to the Nobel Prize and Beyond*, Nobel Lecture, December 2006.

[24] K. D. Miller, $\pi =$ *Visual Cortex*, Science **330** (2010): 1059-60.

[25] L. Nicolaescu, *Critical sets of random smooth functions on compact manifolds.* arXiv preprint (2010): arXiv:1008.5085

[26] J. F. Nye & M. V. Berry, *Dislocations in wave trains*, Proc. R. Soc. Series A **336** (1974), 165-90

[27] G. S. Xia, S. Ferradans, G. Peyré & J. F. Aujol, *Synthesizing and mixing stationary Gaussian texture models.* SIAM Journal on Imaging Sciences, **7(1)** (2014) 476-508.

[28] G. Smoot, *Cosmic Microwave Background Radiation Anisotropies: their Discovery and Utilization*, Nobel Lecture, December 2006.

[29] E. F. G. Thomas, *An infinitesimal characterization of Gel'fand pairs*, Contemp. Math. **26** (1984), 379-385.

[30] N. Ja. Vilenkin & A. U. Klimyk, *Representation of Lie Groups and Special Functions* (3 volumes), Mathematics and Its Applications (Soviet Series), 1991

[31] L. Y. Wei, S. Lefebvre, V. Kwatra & G. Turk, *State of the Art in Example-based Texture Synthesis*, Eurographics, 2009

[32] F. Wolf and *T. Geisel*, *Spontaneous pinwheel annihilation during visual development*, Nature **395** (1998), 73-78.

[33] F. Wolf and T. Geisel, *Universality in visual cortical pattern formation.* . Journal of Physiology-Paris **97(2)** (2003) 253-264.

[34] J. A. Wolf, *Harmonic Analysis on Commutative Spaces*, Mathematical Surveys and Monographs 142, AMS, 2007.

[35] J. A. Wolf, *Spaces of Constant Curvature* (sixth edition); AMS Chelsea publishing,

[36] A. M. Yaglom, *Second-order homogeneous random fields*. Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, Volume 2: Contributions to Probability Theory. The Regents of the University of California, 1961.

[37] A. Zerva, *Spatial variation of seismic ground motions*, Advances in Engineering series, CRC Press (2009)

Part II

Does the vestibulo-cerebellum work with unitary representations of the Galilei group ?

Chapter 5

Representations of the Galilei group

Contents

Abstract. This chapter gathers some facts on the Galilei group and its unitary or projective representations. I discuss the relationship between the Schrödinger equation and the projective representations of the Galilei group, recall some of Inönü and Wigner's work on the unitary irreducible representations, and for the purposes of the next chapter, record formulae (due to Vilenkin and Miller) for the matrix elements of some of the unitary irreducible representations.

The examples here considered provide a good opportunity to state Mackey's results on the representation theory of semidirect products, which are the foundation for Part III, and to describe a somewhat simplified proof for these results in the cases considered in this thesis.

I said in the introduction that the Galilei group should be particularly appropriate to discuss the vestibular system and try to understand how it senses and plans our motions. I also recalled that when Inönü and Wigner introduced Lie group contractions, it was this group, and the relationship between the relativistic and nonrelativistic versions of quantum mechanics, which they had in mind. I alluded to specific features of the Galilei group already in Chapter 4, and I will use some of its unitary representations as my main tool in the next chapter. But this group is less famous than might be expected given its historical and conceptual importance. As a consequence, the present chapter gathers some of the key facts about the Galilei group and some of its representations.

Since the Galilei group is a semidirect product with a normal vector subgroup, a discussion of its representations is a good opportunity state and prove a version of Mackey's theorem on the representations of semidirect products of this kind.

What I am going to recall below is quite well-known. Most of sections 1, 2 and 3.1 (and much more on the physical side) is discussed by Lévy Leblond in [1] ; his paper is to my knowledge the most comprehensive survey on the Galilei group. Many fascinating facts about the group can also be found in Souriau's book [8], though the emphasis there is put on classical mechanics and geometric quantization rather than representation theory.

1 The definition; some structure properties

To introduce the Galilei group, it is rather natural to start with the physical notion of "reference frame". One task we can attribute, at least in principle, to every "observer", is that of assigning a position and a date to physical events. In mathematical-sounding (but not mathematically sound) language, it is perhaps appropriate to imagine that there "is" an ill -defined "set $\mathcal{E}vents$ of events" and that with each choice of reference frame there comes a "map" $\mathcal{E}vents \to \mathbb{R}^3 \times \mathbb{R}$.

Galilean relativity is based on the idea that observers should agree on:

(a) the duration (with sign) between events,

(b) the (oriented) shape of solid objects,

(c) the notion that a solid object has a uniform translational motion.

Without that agreement the notions of objective duration, solid object, and uniform motion lose their meaning, and indeed the fact that (b) is not satisfied means that the notion of solid object is stripped of its meaning in relativistic mechanics.

Now, conditions (a), (b) and (c) are strong enough to determine the relationship between the "maps" $\mathcal{O} : \mathcal{E}vents \to \mathbb{R}^3 \times \mathbb{R}$ used by observers who agree on the three notions. Suppose there is a bijection $(x,t) \overset{\mathcal{change}}{\longrightarrow} (\mathbf{p}(x,t), \mathbf{d}(x,t))$ which can be used to translate one $\mathcal{O}$-map into another by composition. Then

(a) means that $\mathbf{d}(x,t)$ must be equal to $\mathbf{d}(x,0)+t$ for every x and t ;

(b) means that for every t, the map $x \mapsto \mathbf{p}(x,t)$ must be an (orientation-preserving) affine isometry,

(c) means that $\mathcal{change}$ must send aligned points of $\mathbb{R}^4$ to aligned points, thus be affine.

The Galilei group gathers the transformations which have these three properties. It is thus the subgroup of the affine group of $\mathbb{R}^4$ which gathers the maps

$$(x,t) \mapsto (Ax + vt + x_0, t + t_0) \quad \text{with} \quad A \in SO(3), v \in \mathbb{R}^3, x_0 \in \mathbb{R}^3, t_0 \in \mathbb{R}.$$

I will write $\mathbf{G}$ for the Galilei group, and when (A, v, x_0, t) is an element of $SO(3) \times \mathbb{R}^3 \times \mathbb{R}^3 \times \mathbb{R}$, I will write $\mathcal{Gal}(A, v, x_0, t_0)$ for the above map.

Using homogeneous coordinates to turn an affine transformation of $\mathbb{R}^5$ into a linear transformation of $\mathbb{R}^5$, we obtain a faithful representation of the Galilei group: $\mathbf{G}$ is isomorphic with the subgroup

$$\left\{ \begin{pmatrix} A & v & x \\ 0 & 1 & t \\ 0 & 0 & 1 \end{pmatrix} \mid A \in SO(3), v \in \mathbb{R}^3, x \in \mathbb{R}^3, t \in \mathbb{R} \right\}$$

of $GL_5(\mathbb{R})$.

It is perhaps useful to record in print that

$$\mathcal{Gal}(A_1, v_1, x_1, t_1)\mathcal{Gal}(A_2, v_2, x_2, t_2) = \mathcal{Gal}(A_1 A_2 \ , \ v_1 + A_1 v_2 \ , \ x_1 + Ax_2 + v_1 t_2 \ , \ t_1 + t_2);$$
$$\mathcal{Gal}(A, v, x_0, t)^{-1} = \mathcal{Gal}\left(A^{-1}, -A^{-1}v, A^{-1}(x_0 - vt_0), -t_0\right). \tag{1.1}$$

The *homogeneous Galilei group* is the subgroup $\mathbf{G} \cap GL_4(\mathbb{R}) = \{\mathcal{Gal}(A, v, 0, 0) \mid (A, v) \in SO(3) \times \mathbb{R}^3\}$. I will write $\mathbf{G}_{hom}$ for it. It is isomorphic with the Euclidean group of affine, orientation-preserving isometries of $\mathbb{R}^3$. It is $\mathbf{G}_{hom}$, rather than $\mathbf{G}$, which will be used in the next chapter to analyze the activity of vestibular neurons in the cerebellum.

Crucial for the determination of the unitary representations of $\mathbf{G}$ is the fact that the space-time translation subgroup

$$\mathbf{E} = \left\{ \mathcal{Gal}(1, 0, x, t) \mid (x, t) \in \mathbb{R}^3 \times \mathbb{R} \right\}$$

is abelian and normal in $\mathbf{G}$. The Galilei group in fact is the semidirect product

$$\mathbf{G} = \mathbf{G}_{hom} \ltimes \mathbf{E}$$

associated to the action $(A, v) \cdot (x, t) \mapsto (Ax + vt, t)$ of $\mathbf{G}_{hom}$ on $\mathbf{E}$. Thus the homogeneous Galilei group $\mathbf{G}_{hom}$, in addition to being a subgroup of $\mathbf{G}$, also appears as the quotient $\mathbf{G}/\mathbf{E}$.

2 Mass as a cohomology class, Projective representations, and the Schrödinger equation.

I recalled in the introduction that Wigner's description of the irreducible unitary representations of the Poincaré group was a breakthrough of the first importance for physics, opening the way to a description (or rather, definition) of elementary particles. Remarking that a unitary representation is the same thing as a projective representation in the Poincaré case (I will recall in this section why that is the case) and that the conceptual structure of quantum mechanics made the study of projective representations more natural than the study of unitary ones, Bargmann studied the projective representations of the Galilei group in 1954, and proved that the Schrödinger equation is that determining the carrier space for a generic projective representation of $\mathbf{G}$ whose "little-group parameter" (to be defined below) is trivial.

From section 3 onwards (and in the next chapters), the focus will be on "true" (linear) representations rather than projective ones. However, enthusiastic reactions to Bargmann's results from young colleagues seem to indicate that after more than sixty years, the link between Schrödinger's equation and group theory is worth popularizing again. The argument boils down to the easiest part Mackey's "little group" method which is crucial to the second part of my thesis, and it is desirable that I give at least a sketch of proof for the method. So I shall pause here to describe Bargmann's 1954 theorem. The contents of this subsection are quite well-known − my only reason for including them to this chapter is that I have been unable to find an exposition which does not assume previous knowledge of Schrödinger's equation, but instead exhibits it directly from the group structure. I should recall clearly that in the cerebellum-related discussions of the next chapter, I am going to emphasize unitary representations rather than projective ones; upon describing the unitary representations in Section 3, I am also going to recall why unitary representations quickly dropped out of focus in physics.

2.1 Projective representations and the first cohomology group

Suppose G is a *connected and simply connected* Lie group. Then a map $T : G \to \mathcal{U}(\mathbf{H})$ is[1] a projective representation of G if there is a function $\xi : G \times G \to \mathbb{R}$ such that

$$T(g_1 g_2) = e^{i\xi(g_1,g_2)} T(g_1) T(g_2).$$

In that case, $\xi(1_G, 1_G)$ is a multiple of 2π, and the function ξ is determined uniquely by T if we ask that $\xi(1_G, 1_G)$ be zero. Because group multiplication is associative, we then have

$$\xi(g_1, g_2) + \xi(g_1 g_2, g_3) = \xi(g_1, g_2 g_3) + \xi(g_2, g_3) \tag{2.1}$$

for g_1, g_2, g_3 in G.

A function $\xi : G \times G \to \mathbb{R}$ for which (2.1) is satisfied is called a G*-cocycle.* An important example is furnished by functions ξ which read

$$\xi_\zeta(g_1, g_2) = \zeta(g_1 g_2) - \zeta(g_1) - \zeta(g_2). \tag{2.2}$$

with $\zeta : G \to \mathbb{R}$ and $\zeta(1_G) = 0$. Such ξs are of course G-cocycles, and they are special among G-cocycles because if such a cocycle comes with a projective representation T as above, then

$$g \mapsto e^{i\zeta(g)} T(g) \tag{2.3}$$

is immediately seen to be unitary. Such cocycles are called G*-coboundaries.* To explain the cohomological terminology, let us note that both (2.1) and (2.2) are stable under linear combinations with real coefficients, and form the vector space quotient

$$H^1(G) := \{\ G\text{-cocycles}\ \} \,/\, \{\ G\text{-coboundaries}\ \}.$$

When $H^1(G)$ is trivial, every projective representation can be made unitary in the way indicated by (2.3). In general, finding a representative for each class in $H^1(G)$ makes

1. We take this as a definition for convenience and brevity, but projective representations can of course be defined for disconnected or non-simply-connected groups, and the definition of (global or local) G-cocycles has to be adapted; see Bargmann's 1954 paper.

it possible to reduce the study of projective representations to that of unitary representations: if T_1 and T_2 are two projective representations which determine two G-cocycles whose difference is a G-coboundary ξ_ζ, then $g \mapsto e^{i\zeta(g)}T_2(g)$ and T_1 determine the same G-cocycle; as we shall see, the study of projective representations with a given associated G-cocycle, say ξ, then reduces to study to that of unitary representations of a central extension of G determined by ξ.

When Bargmann said all this in 1954, he accompanied the discussion with a determination of the group $H^1(G)$ for several of the groups that were then of interest in physics.

Theorem (Bargmann, 1954).

1. *Suppose G is a connected and simply connected semisimple Lie group. Then $H^1(G)$ is zero.*
2. *Suppose L is a connected and simply connected semisimple Lie group, V is a vector space equipped with a linear action of L, and G is the semidirect product $L \ltimes V$ built from that linear action. Suppose $\dim V \geq 3$. Then $H^1(G) = 0$.*
3. *Suppose $\tilde{\mathbf{G}}$ is the universal covering of the Galilei group. Then $H^1(\tilde{\mathbf{G}})$ is one-dimensional.*

The relationship between the Galilei group and its universal covering is particularly simple. Recall that there is a group morphism from the connected and simply connected group $SU(2)$ to $SO(3)$ with a two-element kernel. This yields an action of $SU(2)$ on $\mathbb{R}^3$; the corresponding semidirect product is the universal covering of $\mathbf{G}_{hom}$, which in turn acts on E. A group law can then be defined on $SU(2) \times \mathbb{R}^3 \times \mathbb{R}^3 \times \mathbb{R}$ by imitating (1.1), and I shall write $\tilde{\mathbf{G}}$ for the corresponding group. I shall thus extend the notation $\mathcal{Gal}(A, v, x, t)$ to the case in which A is in $SU(2)$. The obvious map from $\tilde{\mathbf{G}}$ to $\mathbf{G}$ is then a two-fold covering.
Bargmann's calculation of $H^1(\tilde{\mathbf{G}})$ can be made explicit as follows: if $g_0 = \mathcal{Gal}(A_0, v_0, x_0, t_0)$ and $g_1 = \mathcal{Gal}(A_1, v_1, x_1, t_1)$ are elements of $\tilde{\mathbf{G}}$, define

$$\xi_m(g_1, g_0) = m\left(\frac{1}{2}t_1 \|v_0\|^2 + v_1 \cdot (A_1 x_0)\right).$$

Then ξ_m is a $\tilde{\mathbf{G}}$-cocycle, and it is not a $\tilde{\mathbf{G}}$-coboundary unless m is zero; as m ranges over $\mathbb{R}$, the corresponding cohomology classes exhaust $\mathbf{H}^1(\tilde{\mathbf{G}})$.

⋆

One of the most important differences between Galilean and special relativity is the status of mass in both physical theories. This important fact alluded to in Chapter 5 can be viewed as a physical reflection of the difference between the first cohomology group of the Poincaré group and that of Galilei group; since the specifically relativistic character of the "mass-energy equivalence", through some of its practical consequences, is among the most famous facts of all science, it is perhaps appropriate to see how the above considerations on group cohomology can help discuss it. One way to do that would be to start with Schrödinger's equation, to show that the mass which appears there determines a nonzero element in $H^1(\mathbf{G})$, and to compare this with the status of mass in, say, the Klein-Gordon or Dirac equations. This is certainly relevant for quantum mechanics, and this is the way Bargmann showed that the Schrödinger equations corresponding to nonzero masses exhaust the projective representations which determine nonzero classes in $H^1(\mathbf{G})$. As I said, I will take things in the reverse order and show how to *obtain* the Schrödinger equation directly from group theory, without assuming previous knowledge of it..

2.2 The irreducible projective representations of the Galilei group and Schrödinger's equation

The aim of this paragraph is to show that Schrödinger's equation for the wave function ψ of a single free particle of mass m,

$$\frac{\partial \psi}{\partial t} + \frac{i\hbar}{2m}\Delta\psi = 0,$$

can be obtained from the requirement that its solution space be the carrier space for a projective irreducible representation of $\mathbf{G}$ whose cohomology class is that of ξ_m.

We first recall how the study of projective representations whose associated $\tilde{\mathbf{G}}$-cocycle is ξ_m can be reduced to the study of unitary representations of a central extension of $\tilde{\mathbf{G}}$. We can define a new group $\tilde{\mathbf{G}}_m$ as the set $\tilde{\mathbf{G}} \times \mathbb{R}$, together with the composition

$$(g_1, \alpha_1) \cdot (g_0, \alpha_0) := (g_1 g_0, \alpha_1 + \alpha_0 + \xi_m(g_1, g_0))$$

for g_1, g_0 in $\tilde{\mathbf{G}}$ and α_1, α_2 in $\mathbb{R}$. Equation (2.1), aside from the fact that $\tilde{\mathbf{G}}_m$ is indeed a group, shows that $\{0\} \times \mathbb{R}$ is in the center of $\tilde{\mathbf{G}}_m$.

An immediate consequence of the fact that the space-time translation subgroup $\mathbf{E}$ is abelian and normal in $\tilde{\mathbf{G}}$ is that $\mathbf{E} \times \mathbb{R}$ is abelian and normal in $\tilde{\mathbf{G}}_m$. So the extended group $\tilde{\mathbf{G}}_m$ appears as a semidirect product with a five-dimensional abelian normal factor:

$$\tilde{\mathbf{G}}_m = \tilde{\mathbf{G}}_{hom} \ltimes (\mathbf{E} \times \mathbb{R}).$$

Its unitary representations may then be determined through Mackey's "little group" method.

The method is of the first importance for all the parts of this thesis that remain to be discussed; for our purposes in this section it will be enough to describe some of the beginnings of the proof here (a description of Mackey's method will be completed in the next section with a more complete statement and proof).

⋆

Suppose A is a finite-dimensional vector space, H is a unimodular Lie group, $\rho : H \to GL(A)$ is a group morphism, and $G = H \ltimes_\rho A$ is the corresponding semi-direct product; for simplicity I shall assume [2] that the action of H on A is volume-preserving. In that case G, and the subgroups $H' \ltimes A$ obtained from closed unimodular subgroups H' of H are unimodular, and there will be pleasant simplifications in the formulae to come.

I will describe how one associates the closure of a G-orbit in $\widehat{A}$ to every irreducible representation of G, at least when every H-orbit in $\widehat{A}$ is locally closed. In several interesting cases, this procedure will in fact yield a unique G-orbit.

Suppose π is a unitary representation of G on a Hilbert space $\mathcal{H}$. Set

$$P_\pi := \left\{ \text{continuous functions } \varphi \text{ on } \widehat{A}, \text{ vanishing at infinity, such that } \int_A \widehat{\varphi}(x)\pi(x)dx = 0 \right\}$$

(the integral is well-defined for such φ because the Fourier transform $\widehat{\varphi}$ is integrable on A). This is a closed ideal in $\mathcal{C}_0(\widehat{A})$; define $\mathcal{C}_\pi \subset \widehat{A}$ as its zero-set:

$$\mathcal{C}_\pi := \left\{ \chi \in \widehat{A} : \ \varphi(\chi) = 0 \text{ for all } \varphi \text{ in } P_\pi \right\}.$$

2. This is not a hypothesis, but a fact, for each of the semidirect products considered in this thesis.

Then $\mathcal{C}_\pi$ is a closed G-invariant subset of $\widehat{A}$. Since P_π is a closed ideal, it coincides with the space of functions which vanish on $\mathcal{C}_\pi$.

A key observation is now that *if π is irreducible, there turns out to be a H-orbit Ω in $\widehat{A}$ of which $\mathcal{C}_\pi$ is the closure.*

Indeed, suppose $\mathfrak{U}$ is a countable basis for the topology of $\mathcal{C}_\pi$, and arrange for the empty set not to be in that basis. Start with U in $\mathfrak{U}$. Suppose the closure V_1 of $G \cdot U$ is not all of $\mathcal{C}_\pi$, and set $V_2 = \mathcal{C}_\pi - G \cdot U$. Then V_1 and V_2 are disjoint closed G-invariant subsets of $\mathcal{C}_\pi$. For $i = 1, 2$, the closed subspace $\mathcal{H}_i$ of $\mathcal{H}$ generated by

$$\left\{ \int_A \widehat{\varphi}(x)\pi(x)v dx \mid \varphi \in C_0(\widehat{A}),\ \varphi|_{V_i} \equiv 0,\ v \in \mathcal{H} \right\}$$

cannot be zero because if it were, V_i would be equal to $\mathcal{C}_\pi$. Since π is irreducible, we get $\mathcal{H}_1 = \mathcal{H}_2 = \mathcal{H}$. But this is impossible: if $\mathcal{H}_2$ were to be $\mathcal{H}$, then $\mathcal{V} = \left\{ \int_A \widehat{\varphi}(x)\pi(x)\beta dx \mid \varphi \in C_0(\widehat{A}),\ \varphi|_{V_1} \equiv 0,\ \beta \in \mathcal{H}_2 \right\}$ would be equal to $\mathcal{H}_1 = \mathcal{H}$, but substituting the definition of $\mathcal{H}_2$ into that of $\mathcal{V}$, an elementary calculation using the properties of the Euclidean Fourier transform shows that $\mathcal{V}$ actually must be $\{0\}$.

As a result, the closure $G \cdot U$ *is* all of $\mathcal{C}_\pi$. Then the intersection $\bigcap_{U \in \mathfrak{U}} G \cdot U$ is nonempty because of the Baire theorem. Suppose Ω is the G-orbit of some element in the intersection. Then Ω must meet every open set in the basis $\mathfrak{U}$, so it must be dense in $\mathcal{C}_\pi$. This completes the proof of the observation.

⋆

I have thus shown that every irreducible unitary representation π of G determines a closed subset $\mathcal{C}_\pi$ in which there is a dense G-orbit. It often [3] happens that $\mathcal{C}_\pi$ is itself a closed G-orbit, say Ω.

In the reverse direction, there is a very simple way of associating an irreducible unitary representation of G to every closed orbit Ω in $\widehat{A}$. Consider

$$\left\{ f \in C_c(A)\ ,\ \mathrm{Supp}\widehat{f} \subset \Omega \right\}.$$

Because the Fourier transform turns a translational shift of the variable into multiplication by a (nonconstant) nonvanishing function, and because the Fourier transform is H-equivariant thanks to our volume-preservation hypothesis, the natural (quasi-regular) action of G on functions on A leaves that space invariant. Its completion with respect to the usual $\mathbf{L}^2$ inner product thus carries a unitary representation of G. Since Ω has no proper G-invariant subset, the representation does look irreducible. The formal statement that this is the case is part of Mackey's work on induced representations and systems of imprimitivity, because our irreducible representation is immediately seen to be equivalent with that induced from the trivial representation of the stabilizer of any point of Ω under the action of G on $\widehat{A}$ (see the next section).

⋆

Of course $\widehat{A}$ is a finite-dimensional vector space. In favourable cases, it is not difficult to write down a wave equation whose solution space is the above-mentioned representation :

3. In this thesis, we meet semidirect products $K \ltimes A$ with K compact, the Galilei group and the Poincaré group: for those groups, the only cases in which $\mathcal{C}_\pi$ is not a single G-orbit are the zero-mass (light-like) representations of the Poincaré group which are not trivial upon restriction to the space-time translation subgroup.

when Ω is an affine algebraic subset of $\widehat{A}$, the condition

$$\text{Supp } \widehat{f} \subset \Omega$$

is a linear partial differential equation on f. We then obtain a recipe for producing "elementary" linear partial differential equations associated to the algebraic structure of G.

Let me now indicate this symmetry-based recipe produces the Schrödinger equation when it is applied for the Galilei group. As a preliminary remark, let me note that in the study of projective representations of a group G, going up to the universal covering $\tilde{G}$ is natural, but that when G is a semidirect product $H \ltimes A$ as above, the procedure just described automatically produces representations of the universal covering $\tilde{G} = \tilde{H} \ltimes A$ which factor through the covering morphism $\tilde{G} \to G$. So let us start with the extended group $\tilde{\mathbf{G}}_m$ and determine the $\tilde{\mathbf{G}}_{hom}$-orbits in $\widehat{\mathbf{E} \times \mathbb{R}}$: we shall obtain unitary representations of $\mathbf{G}_m$, hence projective representations of $\mathbf{G}$.

Let me write an element of $\widehat{\mathbf{E} \times \mathbb{R}}$ as $(x, t, \alpha) \mapsto e^{i(\langle p, x\rangle + Et + \eta\alpha)}$ with p in $\mathbb{R}^3$, E in $\mathbb{R}$ and η in $\mathbb{R}$. Then the action of $\tilde{\mathbf{G}}_{\mathbf{hom}}$ on $\widehat{\mathbf{E} \times \mathbb{R}}$ is

$$\mathcal{G}al(A, v) \cdot (p, E, \eta) = (Ap + \eta m v, E + \langle v, Ap\rangle + \frac{1}{2}\eta m v^2, \eta).$$

From this formula, it is quite clear that the quantity

$$\eta E - \frac{\|p\|^2}{2m}$$

is constant along every $\mathbf{G}_{hom}$-orbit. Choose U in $\mathbb{R}$; then for every nonzero η,

$$\Omega_{U,\eta} := \left\{ (p, E, \eta) \mid E - \frac{\|p\|^2}{2m \cdot \eta} = U \right\}$$

is in fact a single G-orbit (it is the orbit of $(0, \eta \cdot U, \eta)$; note that the orbit of $(0, U, 0)$ is a point). The above discussion says that an irreducible representation of $\mathbf{G_m}$ is obtained by considering smooth (tempered) functions on $\mathbf{E} \times \mathbb{R}$ whose Fourier transform are supported on $\Omega_{U,\eta}$. A function of this kind reads $(x, t, \alpha) \mapsto e^{i\eta\alpha}\psi(x, t)$, where ψ is a smooth function on $\mathbb{R}^4$ which satisfies the condition

$$(E - \frac{\|p\|^2}{2m \cdot \eta} - U)\widehat{\psi} = 0. \tag{2.4}$$

The dependence on α is completely determined by η and can be forgotten so as to obtain a projective representation realized on a space of functions on $\mathbf{E}$ (the "wave functions"). But (2.4) just means that these functions must be solutions of

$$\frac{\partial \psi}{\partial t} + \frac{i\hbar}{2m \cdot \eta}\Delta\psi = iU.\psi \tag{2.5}$$

When $U = 0$ and $\eta = 1$, (2.5) is of course the Schrödinger equation for a particle of mass m. When U is nonzero and ψ is a solution of (2.5), switching to the map $(x, t) \overset{\tilde{\psi}}{\mapsto} e^{iUt}\psi(x, t)$ turns a solution of (2.5) into a solution of the Schrödinger equation (with either m or $\hbar$ rescaled), and at this point it might not be out of place to recall that if $x \mapsto \psi(x, t)$ is to represent a probability amplitude whose squared modulus describes the probability distribution for the position of an isolated particle at t, $\tilde{\psi}$ and ψ are not physically distinct.

3 Unitary irreducible representations: the work by Inönü and Wigner and explicit formulae for some matrix elements

The significance of Bargmann's work for physics is quite clear; now, although the focus in quantum mechanics is understandably on projective representations rather than unitary ones, it is quite natural to wonder whether the unitary representations of the Galilei group do have physical significance: in the interval between Bargmann's discovery and its appearance in print in 1954, Wigner and Inönü (who had been in contact with Bargmann) published their work on the unitary representations of **G**. Here are hints to their 1952 solution (the original paper is [3]; it also seems to mark the first appearance Galilei group in print. For later − but easier to read and mathematically sounder − discussions of the solution, see [4] and [1]) .

3.1 Inönü and Wigner's parameters for the unitary representations of G, and Mackey's "little group" theorem.

We earlier built the Hilbert space for an irreducible representation of G as a space of functions on A whose Fourier transform is concentrated on a single G-orbit in $\widehat{A}$. Other Hilbert spaces can be obtained by “twisting" that construction to produce spaces of sections of vector bundles, as follows. Start with χ in $\widehat{A}$, let me write H_χ for the stabilizer (widely known in physics as the *little group at* χ) of χ for the action of H on $\widehat{A}$ and G_χ for $H_\chi \ltimes A$. Suppose $(\mathfrak{H}, \mu)$ is an irreducible representation of H_χ. The full group G acts on the the homogeneous vector bundle $G \times_{G_\chi} \mathfrak{H}$ (viewed as a vector bundle over the orbit of χ with fiber $\mathfrak{H}$) that can be built from μ, it acts on its sections too, so that a suitable completion of the space of compactly supported sections [4] produces a unitary representation of G. I shall now recall that Mackey proved that it is irreducible, and that these constructions exhaust the unitary dual of G. Mackey's theory of induced representations came to full maturity in the late 1950s, but the special case we need actually predates Inönü and Wigner's work [5]. Mackey's original paper is [5]; see also [6], [7].

Mackey's formulation is in terms of induced representations. Suppose $G = H \ltimes A$ is a semidirect product as above. Suppose χ is an element in the vector space dual $A^\star$, and $(\mathfrak{H}, \mu)$ is an irreducible H_χ-module (I shall call (χ, μ) a *Mackey datum*). Instead of considering the associated vector bundle, define a representation of G_χ as $\tilde{\mu} = \mu \otimes \chi$, consider

$$\mathcal{V}_{\chi,\mu} := \Big\{ f : G \xrightarrow{C_c} \mathfrak{H} \ : \ f(g\gamma) = \tilde{\mu}(\gamma) f(g) \ , \text{ for all } \gamma \text{ in } G_\chi \Big\} \tag{3.1}$$

and complete it with respect to the $\mathbf{L}^2$ inner product. Write $\mathcal{H}_{\chi,\mu}$ for the completion, and equip it with the action of G which extends that on $\mathcal{V}_{\chi,\mu}$ in which g acts through $f \mapsto (x \mapsto f(g \cdot x))$.

Mackey's 1949 theorem.

(a) *If* (χ, μ) *is a Mackey datum, the representation of* G *on* $\mathcal{H}_{\chi,\mu}$ *described above is irreducible.*

(b) *If* (χ_1, μ_1) *and* (χ_2, μ_2) *are Mackey data, a necessary and sufficient condition for the irreducible representations of* G *on* $\mathcal{H}_{\chi_1,\mu_1}$ *and* $\mathcal{H}_{\chi_2,\mu_2}$ *to be equivalent is that there be*

4. In the case, not considered here, of a non-unimodular G, several adjustments are necessary.

an element of H which sends χ_1 to χ_2, say h, and that $\mu_1 \circ h^\star$ be equivalent with μ_2 as a H_{χ_2}-module.

(c) *Every unitary irreducible representation of G is equivalent with one built from a Mackey datum.*

Mackey data in which the little-group representation μ is trivial and the orbit of χ is closed give rise to the representations of G discussed above; that is why I said in the introduction to this chapter that the Schrödinger equation is naturally tied to projective representations of **G** with "little-group parameter" trivial.

⋆

Fundamental as it is, Mackey obtained his result as a consequence of a much more general theorem valid for locally compact groups, and his proof takes many subtle measure-theoretic issues into account. I am going to include here a summary of a shorter proof which is due to Orsted [15] and is based on ideas by Blattner [14] and especially Poulsen [16]. Every single detail is checked in Chapters 3 and 4 of Kaniuth and Taylor's recent book [17].

The main conceptual point in the theorem is part (c), so I shall first focus on proving (c). Let us start with a unitary irreducible representation π of G on a Hilbert space $\mathcal{H}$. Recall that we built a closed subset $\mathcal{C}_\pi$ of $\widehat{A}$ above, and that there is a single G-orbit of which it is the closure. I shall now assume, mainly for simplicity, that it *is* a single orbit, say Ω. As I said this is automatic for each of the semidirect products whose unitary dual is considered in detail in this thesis.

Choose χ on Ω ; I will start by associating a bounded operator on $\mathcal{H}$ to every element of $\mathcal{C}_c\left(G/G_\chi\right)$. Beginning with φ in $\mathcal{C}_c\left(G/G_\chi\right)$, we can view it as an element of $\mathcal{C}_c(\Omega)$, and set

$$E(\varphi) := \int_A \widehat{\Phi}(a)\pi(a)da$$

as soon as Φ is an element of $\mathcal{C}_c(\widehat{A})$ whose restriction to Ω is φ. This can be done because if Φ_1 and Φ_2 are two such elements, then $(\Phi_1 - \Phi_2)$ vanishes on Ω ; because the space we called P_π above is a closed ideal in $\mathcal{C}_0(\widehat{A})$, it is determined by the common-zero set of its elements, so $(\Phi_1 - \Phi_2)$ actually lies in it and thus $\int_A \widehat{\Phi_1}(a)\pi(a)da = \int_A \widehat{\Phi_2}(a)\pi(a)da$.

The map E sends $\mathcal{C}_c\left(G/G_\chi\right)$ to the bounded operators on $\mathcal{H}$; in Mackey's terminology E and π determine a *system of imprimitivity for G based on G/G_χ*.

To prove Mackey's theorem I need to use π to build a representation of G_χ, and the challenge is to carve a carrier space out of $\mathcal{H}$.

Let us start with the operation of averaging a function on G over the right G_χ-cosets: if φ is in $\mathcal{C}_c(G)$, set

$$Av[\varphi] = xG_\chi \mapsto \int_{G_\chi} (x\gamma)d\gamma.$$

This is an element of $\mathcal{C}_c(G/G_\chi)$. Now if x and y are two elements of $\mathcal{H}$, then the map

$$\mu_{x,y} = \varphi \mapsto \langle E\left(Av[\varphi]\right)x\ ,\ y\rangle$$

is continuous with respect to the natural Fréchet topology on $\mathcal{C}_0(G)$, and thus it defines a Radon measure on G. If x and y lie in the Gårding subspace

$$\mathcal{H}^\infty := \operatorname{Span}\left\{\int_G \varphi(g)\pi(g)udg \middle| u \in \mathcal{H}, \varphi \in \mathcal{C}_c(G)\right\},$$

a key observation by Poulsen and Orsted is now that $\mu_{x,y}$ is absolutely continuous with respect to the (two-sided) Haar measure on G, so that it is given by integration against a *continuous* function, say [5] $g \mapsto R_{x,y}(g)$.

This makes it possible to evaluate the density at 1_G, setting $\beta(x,y) = R_{x,y}(1_G)$, and then the map β defines a sesquilinear form on $\mathcal{H}^\infty \times \mathcal{H}^\infty$. A close look at the formulae for $R_{x,y}$ reveals that this sesquilinear form is G_χ-invariant, and that $\beta(x,x) \geq 0$ for every x in $\mathcal{H}^\infty$.

Quotienting out the kernel $\mathcal{K} = \{x \in \mathcal{H}^\infty \ : \ \beta(x,y) = 0 \text{ for all } y \text{ in } \mathcal{H}^\infty\}$, we can reach the aim : we set

$$\mathfrak{H} = \text{ completion of } \mathcal{H}^\infty/\mathcal{K} \text{ w.r.t. the inner product inherited from } \beta,$$

and note that thanks to the invariance properties of β, the G_χ-action on $\mathcal{H}^\infty$ defines a unitary representation $\tilde{\mu}$ of G_χ. Of course G_χ is the semidirect product $H_\chi \ltimes A$.

We have thus obtained from π a couple (χ, μ) which is very likely to correspond to a Mackey datum (we will be sure that is is really a Mackey datum when it will be established that $\tilde{\mu}$ is equivalent with $\mu \otimes \chi$ where μ is some irreducible representation of H_χ) . To show that π is actually equivalent with the representation induced from χ and $\tilde{\mu}$, we note that a function from G to $\mathfrak{H}$ may be associated to every ξ in $\mathcal{H}^\infty$: we need only set

$$f_\xi = g \mapsto \text{projection, in } \mathfrak{H}, \text{ of } \pi(g^{-1})\xi.$$

The map $\xi \mapsto f_\xi$ then easily extends to a linear isometry between $\mathcal{H}$ and $\mathcal{H}^{\chi,\tilde{\mu}}$, and intertwines the G-actions on both spaces. A byproduct of this is that $\tilde{\mu}$ is actually irreducible, otherwise π would not be. Now we can assume that π is the representation induced from $\tilde{\mu}$, and then a simple calculation shows that $\tilde{\mu}|_A$ is in fact χ, so $\tilde{\mu}$ is equivalent with $\mu \otimes \chi$, where μ a unitary irreducible representation of H_χ. Then (χ, μ) is a Mackey datum, and π is equivalent with the representation built from that Mackey datum. With this the proof of (c) is complete.

To prove (a) and (b), we first note that in view of section 2.2, in the case where every H-orbit in $\widehat{A}$ is closed, Mackey data (χ_1, μ_1) and (χ_2, μ_2) can give rise to equivalent representations only if χ_1 and χ_2 lie on the same H-orbit.

Assume now that χ_1 and χ_2 coincide, write χ for the both of them and $\mathfrak{H}_1$ and $\mathfrak{H}_2$ for the carrier spaces of the representations μ_1 and μ_2 of H_χ; then we observe that there is an isomorphism between $\mathrm{Hom}_{H_\chi}(\mathfrak{H}_1, \mathfrak{H}_2)$ and $\mathrm{Hom}_G(\mathcal{H}_{\chi,\mu_1}, \mathcal{H}_{\chi,\mu_2})$.

In fact, to any ι in $\mathrm{Hom}_{H_\chi}(\mathfrak{H}_1, \mathfrak{H}_2)$ one can assign the map from $\mathcal{V}_{\chi,\mu_2}$ to $\mathcal{V}_{\chi,\mu_2}$ defined by composition with ι, and then extend the assignation to a map $\mathfrak{I}(\iota)$ between $\mathcal{H}_{\chi,\mu_2}$ and $\mathcal{H}_{\chi,\mu_2}$ which is obviously G-equivariant. The map $\mathfrak{I}$ is thus a morphism from $\mathrm{Hom}_{H_\chi}(\mathfrak{H}, \mathfrak{H}')$ to $\mathrm{Hom}_G(\mathcal{H}_{\chi_1,\mu_1}, \mathcal{H}_{\chi_2,\mu_2})$, and to prove that it is an isomorphism the simplest strategy is perhaps to write down the inverse

5. This is easily seen by making $\mu_{x,y}$ more explicit when $x = \int_G \psi_x(g)\pi(g)u_x dg$ and $y = \int_G \psi_y(g)\pi(g)u_y dg$ for some ψ_x, ψ_y in $\mathcal{C}_c(G)$ and some u_x, u_y in $\mathcal{H}$. To be precise, for f in $\mathcal{C}_c(G \times G)$, define a function on $(G/G_\chi) \times G$ as $Av_1(f) = (aG_\chi, b) \mapsto \int_{G_\chi} f(a\gamma, b)d\gamma$; then $f \mapsto \int_G \langle E(Av_1[f])\pi(b^{-1}x), y\rangle$ defines a continuous linear functional on $C_c(G \times G)$, and thus a Radon measure on $G \times G$. Writing $d\lambda_{x,y}$ for it, we then have

$$\mu_{x,y}(\varphi) = \int \varphi(g) M_{x,y}(g) dg$$

where $M_{x,y} = g \mapsto \int_{G\times G} \bar{\psi}_2(a^{-1}g)\psi_1(ba^{-1}g)d\lambda_{x,y}(a,b)$. This is a continuous function on g by the usual integration arguments.

morphism explicitly. Recall that when (χ, μ) is a Mackey datum and $\mathfrak{H}$ is the carrier space for μ, evaluation at the identity defines a map from $\mathcal{V}_{\chi,\mu}$ to $\mathfrak{H}$ and that the Gårding-like subspace $\mathcal{V}^{\infty}_{\chi,\mu} := \text{Span}\left\{\int_G \varphi(g)\pi_{\chi,\mu}(g)udg \middle| u \in \mathcal{V}_{\chi,\mu}, \varphi \in \mathcal{C}_c(G)\right\}$ is dense in $\mathcal{V}_{\chi,\mu}$ (in addition to being a well-known general fact of representation theory, this actually follows from the proof of (c) just written). We can then start with I in $\text{Hom}_{H_\chi}(\mathfrak{H}_1, \mathfrak{H}_2)$, define a linear map ι from the dense subspace $\left\{f(1_G) \mid f \in \mathcal{V}^{\infty}_{\chi,\mu}\right\}$ of $\mathfrak{H}_1$ to $\mathfrak{H}_2$ by setting

$$\iota\left[\left(\sum_k \int_G \varphi_k \pi_{\chi,\mu}(g) u_k dg\right)(1_G)\right] = \left(\sum_k \int_G \varphi_k \pi_{\chi,\mu'}(g)[Iu_k]dg\right)(1_G),$$

and extending ι to a linear map from $\mathfrak{H}_1$ to $\mathfrak{H}_2$, still denoted ι. It is then easily seen that ι is H_χ-equivariant (one needs only go through the definitions and see that its restriction to the elements of the form $\left(\int_G \varphi(g)\pi_{\chi,\mu}(g)udg\right)(1_G)$ is naturally H_χ-equivariant as soon as I is G-equivariant). Thus the assignation $I \mapsto \iota$ defines a morphism from $\text{Hom}_G(\mathcal{H}_{\chi,\mu_1}, \mathcal{H}_{\chi,\mu_2})$ to $\text{Hom}_{H_\chi}(\mathfrak{H}_1, \mathfrak{H}_2)$, and going through the definition it turns out to be the inverse of $\mathfrak{I}$.

Point (b) is then an immediate consequence of the fact that $\text{Hom}_G(\mathcal{H}_{\chi_1,\mu_1}, \mathcal{H}_{\chi_2,\mu_2})$ is zero if and only if $\text{Hom}_{H_\chi}(\mathfrak{H}_1, \mathfrak{H}_2)$ is[6], while point (a) is an immediate consequence of the fact that $\text{Hom}_G(\mathcal{H}_{\chi,\mu}, \mathcal{H}_{\chi,\mu})$ is one-dimensional when $\text{Hom}_{H_\chi}(\mathfrak{H}, \mathfrak{H})$ is. □

⋆

Let me now recall what this means for the unitary irreducible representations of the Galilei group.

The orbits and stabilizers for the action of $\mathbf{G}_{hom}$ in $\widehat{R}^4$ are easily determined: if χ is in $\widehat{\mathbb{R}^4}$ and decomposes as $(x, t) \mapsto \langle p_\chi, x\rangle + E_\chi t$, then $\chi(A^{-1}\vec{x} - (A^{-1}v)t) = \langle Ap_\chi, \vec{x}\rangle + (E_\chi + \langle p_\chi, v\rangle)\, t$ for every A in $SO(3)$ and every v in $\mathbb{R}^3$. This means that if $h = \mathcal{Gal}(A, v, 0, 0)$ is in $\mathbf{G}_{hom}$, then

$$h \cdot \chi := h \cdot (p_\chi, E_\chi) = (Ap_\chi, E_\chi + \langle p_\chi, v\rangle)\,.$$

The orbits of $\mathbf{G}_{hom}$ on $\widehat{\mathbb{R}^4}$ are thus the three-dimensional "cylinders" $\mathcal{C}_\kappa := \{(p, E) \mid \|p\| = \kappa\}$, $\kappa > 0$, and the points $\{(0, E)\}$, $E \in \mathbb{R}$. The stabilizer of $(0, E)$ is of course all of $\mathbf{G}_{hom}$, while the stabilizer of $\chi = (p, E)$ for nonzero p is the subgroup

$$\mathbf{G}_{hom}(\chi) := \{\mathcal{Gal}(A, v; 0; 0) \; : \; Ap = p \text{ and } \; v \perp p\}\,,$$

isomorphic with[7] the Euclidean motion group *of a plane.*

Mackey's theorem yields the following list for the irreducible unitary representations of $\mathbf{G}$.

<u>*Case 1 :*</u> we start with ω in $\mathbb{R}$ and with a unitary irreducible representation U of $\mathbf{G}_{hom}$ on a Hilbert space $\mathcal{H}$, and extend it to the representation $\mathbf{G}$ in which $\mathcal{Gal}(A, v, x, t)$ acts on $\mathcal{H}$ as $e^{i\omega t}U\left[\mathcal{Gal}(A, v)\right]$.

<u>*Case 2:*</u> we start with $\chi = (k, \omega)$ with nonzero k. Because $\mathbf{G}_{hom}(\chi)$ is itself a semidirect product, it representations are obtained by the Mackey method. Since $\mathbf{G}_{hom}(\chi)$ naturally

6. Note that if χ_1 and χ_2 lie on the same H-orbit, writing h for an element of H sending χ_1 to χ_2, the representation built from (χ_2, μ_2) is naturally equivalent with that built from $(\chi_1, \mu_2 \circ h^\star)$, so that we may assume $\chi = \chi'$ when discussing point (b).

7. Of course it directly *is* the Euclidean motion group of $p^\perp$ when the action is the usual action on $\mathbb{R}^3$, but here the action is different as we saw

identifies with the group of Euclidean rigid motions in the plane orthogonal to k, the orbits we need to study are those of the rotations with axis k in the linear space of vectors orthogonal to k. They are circles and the stabilizer of each point on such a circle is trivial, except when the circle has zero radius. The ensuing list of representations of $\mathbf{G}(\chi)$ and $\mathbf{G}$ runs as follows.

Case 2a: we choose a n in $\mathbb{Z}$ and use the one-dimensional representation $\mathcal{G}al(A_\vartheta, v) \overset{\mu_{0,n}}{\mapsto} e^{in\vartheta}$ of $\mathbf{G}(\chi)$. The corresponding representation of $\mathbf{G}$ naturally has a geometric realization on the Hilbert space of square-integrable sections of the n-th tensor bundle of the Hopf bundle on $\mathbf{S}^2$.

Case 2b: we choose a $r > 0$ and use the representation

$$\mathcal{G}al(A_\vartheta, v) \overset{\mu_r}{\mapsto} \left[f \mapsto \left(x \mapsto e^{ir\langle u,v\rangle} f(A_\vartheta^{-1}u)\right)\right]$$

of $\mathbf{G}(\chi)$ on $\mathbf{L}^2(\mathbf{S}^1)$. The corresponding representation of $\mathbf{G}$ naturally has a geometric realization on the Hilbert space of square-integrable sections of the homogeneous vector bundle on $\mathbf{S}^2$ with fiber $\mathbf{L}^2(\mathbf{S}^1)$ that is associated with μ_r.

⋆

As far as quantum physics is concerned, Inönü and Wigner found that the Hilbert spaces obtained in this way are not appropriate as possible homes for particle states. Their argument runs at follows: when transformed into spaces of (vector-valued) functions on $\mathbb{R}^4$, these spaces must be spaces of functions whose Fourier transform is concentrated on a cylinder $\mathcal{C}_\kappa$ or a point. When interpreting these wave functions ψ on $\mathbb{R}^4$ as probability amplitude distributions, the Fourier transform of $x \mapsto \psi(x,t)$ at fixed t is to be interpreted as a probability amplitude distribution for the linear momentum. Because the Fourier transform of a function with small support typically has a large support whereas the space projection of $\mathcal{C}_\kappa$ is compact, Inönü and Wigner argued that none of the above Hilbert spaces can contain vectors representing a particle localized in a small region of space; they went on to add that none can contain any vector representing a particle with zero velocity, and concluded that it is not reasonable to use any unitary irreducible representation space of $\mathbf{G}$ as a receptacle for quantum states.

This explains why the unitary representations of $\mathbf{G}$ seemed less interesting for physics than their projective cousins, or than the unitary representations of the Poincaré group, had proved. There *are* interesting applications of the unitary representations in physics, though; many of them are described in [1].

Our aim in the next chapter, however, was to turn to unitary representations of $\mathbf{G}$ (or rather $\mathbf{G}_{hom}$) as possible providers of special functions which might be of interest in the study of the vestibular system. This calls for a look at matrix elements.

3.2 Explicit formulae for some matrix elements of representations of $\mathbf{G}_{hom}$

In this subsection, I focus on the representations of $\mathbf{G}$ whose associated $\mathbf{G}_{hom}$-orbit is $\{0\}$: these are representations obtained by extending an irreducible representation of $\mathbf{G}_{hom}$ to $\mathbf{G}$ through the projection $\mathbf{G} \to \mathbf{G}_{hom}$, and I am going to write down explicit formulae for their matrix elements in certain bases.

My reason for focusing on the representations of $\mathbf{G}_{hom}$, which are exceptional (and rather trivial) cases in the Mackey machine, is that the subgroup $\mathbf{E}$ is inaccessible to the vestibular sensors of the inner ear; the full Galilei group should be relevant for a system in which the vestibular information is used *together with other senses like vision*. It is $\mathbf{G}_{hom}$, rather than $\mathbf{G}$, which should be of use to discuss the workings of the vestibular system when it is on its own.

Since $\mathbf{G}_{hom}$ is isomorphic with the Euclidean motion group of $\mathbb{R}^3$, what I am going to record below is a set of formulae for "the" matrix elements of each irreducible representation of the Euclidean motion group.

Of course $\mathbf{G}_{hom}$ is itself a semidirect product: writing K for the rotation group $\{\mathcal{G}al(A,0,0,0) \mid A \in SO(3)\}$ and V for the normal abelian subgroup $\{\mathcal{G}al(1,v,0,0) \mid v \in \mathbb{R}^3\}$, $\mathbf{G}_{hom}$ is the semidirect product associated to the action of K on V.

The orbits of K in $\widehat{V}$ can be identified with those in V through the K-invariant usual inner product; they are the spheres centered at the origin.

3.2.a Matrix elements for the irreducible representations of $SU(2)$

Let us start with the single-point orbit situated at the origin; the irreducible representations of $K \simeq SO(3)$ do provide irreducible representations of $\mathbf{G}_{hom}$. Let us record here the result of a calculation by N. Ja. Vilenkin : for the proof see [10], p. 116.

Let ℓ be a nonnegative integer or half-integer. The $(2\ell+1)$-dimensional vector space of all homogeneous polynomials in two complex variables with total degree 2ℓ,

$$\mathfrak{H}_\ell := \left\{ (z_1, z_2) \mapsto \sum_{n=-\ell}^{\ell} a_n z^{n+\ell} z_2^{-n+\ell} \quad , \quad (a_n) \in \mathbb{C}^{2\ell+1} \right\}$$

is [8] stable under the natural action inherited from that of $SU(2)$ on $\mathbb{C}^2$. The corresponding representation is irreducible. Every irreducible representation of $SU(2)$ is equivalent with one $\mathfrak{H}_\lambda$, $\lambda \in \frac{1}{2}\mathbb{N}$.

Vilenkin defines a $SU(2)$-invariant scalar product on $\mathfrak{H}_\ell$ through the requirement that the basis consisting of the monomials $(z_1, z_2) \overset{\mu_n}{\mapsto} z_1^{n+\ell} z_2^{-n+\ell}$, $n = -\ell, ...\ell$, be orthonormal, and for that choice of basis and scalar product, he calculates the matrix elements. After a simple expansion, he finds that if m and n are integers or half-integers of the same kind as ℓ with $|m| \leq \ell$ and $|n| \leq \ell$, the matrix element corresponding to the scalar product between μ_n and the translates of μ_n is

$$t^\ell_{m,n} = \begin{pmatrix} \alpha & \beta \\ -\bar{\beta} & \bar{\alpha} \end{pmatrix} \mapsto \sum_{j=\max(0,n-m)}^{\min(\ell-m,\ell-n)} \binom{\ell-n}{\ell-m-j} \binom{\ell+n}{j} \alpha^{\ell-m-j} \bar{\alpha}^{\ell+n-j} \beta^j (-\bar{\beta})^{\ell+n-j} \tag{3.2}$$

(in the sum the index j is an integer).

3.2.b Matrix elements for the irreducible representations of the Euclidean motion group.

A set of matrix elements for the other representations of $\mathbf{G}_{hom}$ appeared around the same time. What I am going to recall below is due to Willard Miller [12]; his result is described by Chirikjian [13], and the presentation below is based on Chirikjian's paper.

We turn to the representations attached to a sphere $\mathbf{S}$ of radius R in V. The stabilizer of a point p_0 on that sphere is the set K_{p_0} of rotations around a given axis, it is isomorphic

8. In the sum n is an integer or half-integer of the same kind as ℓ

with $SO(2)$; to obtain a representation of $\mathbf{G}_{hom}$ through Mackey's method, we need only (and must) choose a n in $\mathbb{Z}$ and consider the space of sections of the associated vector bundle on $\mathbf{S}$.

We can in fact dispense with nontrivial bundles through a choice of Borel cross-section of the composition $SO(3) = K \to K/K_{p_0} \to \mathbf{S}$. Write $p \mapsto R_p$ for such a cross-section, so that R_p sends p_0 to p when p is on $\mathbf{S}$. For every A in $SO(3)$, the rotation $R_p^{-1} A R_{A^{-1}p}$ is then in K_{p_0}. Call it Σ_A.

Now, equip $\mathbf{L}^2(\mathbf{S})$, the Hilbert space of square-integrable complex-valued functions on $\mathbf{S}$, with a linear action of $\mathbf{G}_{hom}$ by deciding that $\mathcal{G}al(A, v)$ acts on $\mathbf{L}^2(\mathbf{S})$ through

$$f \mapsto \left[p \mapsto e^{i\langle p,v\rangle} \zeta_n\left(\Sigma_A\right) f(A^{-1}p)\right]$$

in which $\zeta_n : K_{p_0} \to \mathbb{C}$ sends the rotation of angle α around p_0 to $e^{in\alpha}$.

Then it is part of Mackey's work on induced representations that the above formula defines a unitary irreducible representation of $\mathbf{G}_{hom}$, one whose equivalence class is that associated with the mackey datum furnished by R and n.

In 1968, Willard Miller chose a suitable basis for $\mathbf{L}^2(\mathbf{S})$ and computed the corresponding matrix elements of the representation under discussion. He started with some generalized Legendre polynomials $P^{\ell}_{m,n}$; these are described in closed form on page 286 in [10] and their relationship with classical, computer-friendly polynomials is discussed there. When $\ell \geq |n|$ and $|m| \leq \ell$ he set

$$h^{\ell}_m(\theta, \phi) = e^{i(m+s)\phi}(-1)^{\ell+s}\sqrt{\frac{2\ell+1}{4\pi}}P^{\ell}_{-s,m}(\cos\theta)$$

for θ in $[0, \pi]$ and ϕ in $[0, 2\pi]$. Viewing θ and ϕ as spherical coordinates, this defines a continuous function on $\mathbf{S}$, and $\left\{h^{\ell}_m\right\}_{\substack{\ell \geq s \\ m=-\ell,\dots\ell}}$ turns out to be an orthonormal basis for $\mathbf{L}^2(\mathbf{S})$.

Miller remarked that for fixed ℓ, the subspace of $\mathbf{L}^2(\mathbf{S})$ spanned by the h^{ℓ}_ms carries an irreducible representation of $SO(3)$, and after comparing the current basis with that chosen in the realization on homogeneous polynomials as above, he found that the restriction to $SO(3)$ of the corresponding matrix elements of $\mathbf{G}_{hom}$ is that we already calculated : using brackets do denote the $\mathbf{L}^2$ scalar product,

$$\langle h^{\ell_1}_{m_1}, \mathcal{G}al(A, 0)h^{\ell_2}_{m_2}\rangle = \delta_{\ell_1,\ell_2} t^{\ell_1}_{m_1,m_2}(A). \tag{3.3}$$

He then proceeded to calculate the restriction of these matrix elements to V. The strategy is to use an expansion of a plane wave as a superposition of spherical harmonics which appears in the physical literature (this seems to be due to Kursunoglu [11], p. 114). Using several "classical" properties of special functions, Miller found that

$$\langle h^{\ell_1}_{m_1}, \mathcal{G}al(1, v)h^{\ell_2}_{m_2}\rangle = \frac{1}{\sqrt{4\pi}} \sum_{\lambda=|\ell_1-\ell_2|}^{\ell_1+\ell_2} i^{\lambda} \sqrt{\frac{(2\lambda+1)(2\ell_1+1)}{2\ell_2+1}} \begin{pmatrix} \lambda & m_2 - m_1 \\ \ell_1 & m_1 \\ \ell_2 & m_2 \end{pmatrix} \begin{pmatrix} \lambda, 0 \\ \ell_1, s \\ \ell_2, s \end{pmatrix} J_{\lambda}(R\,\|v\|)\, Y^{\lambda}_{m_2-m_1}\left(\frac{v}{\|v\|}\right) \tag{3.4}$$

in which the notation $\begin{pmatrix} a & b \\ \alpha & \beta \\ \aleph & \beth \end{pmatrix}$ stands for a Clebsch-Gordan coefficient of $SU(2)$, that which is written as $C(a, b; \alpha, \beta|\aleph, \beth)$ on page 174 of [12], and J_λ stands for the spherical

Bessel function we have been using in Part I.

Now that we know what the restrictions to K and V of our matrix elements are, it is easy to obtain a formula valid on all of $\mathbf{G}_{hom}$: since $\mathcal{G}al(A,v) = \mathcal{G}al(1,v)\mathcal{G}al(A,0)$, we can use (3.3) to get

$$\langle h^{\ell_1}_{m_1}, \mathcal{G}al(A,v) h^{\ell_2}_{m_2}\rangle = \sum_{j=-\ell_2}^{\ell_2} t^{\ell_2}_{j,m_2}(A) \langle h^{\ell_1}_{m_1}, \mathcal{G}al(1,v) h^{\ell_2}_{j}\rangle; \tag{3.5}$$

then formulae (3.2) and (3.4) make the result quite explicit.

⋆

It is perhaps tempting to dismiss (3.2), (3.4) and consequently (3.5) as quite complicated. To resist the temptation, I feel it is appropriate to note that *in these formulae, only elementary functions appear; they are implemented in almost any mathematical software.* In addition, there are well-known recursion formulae for the Clebsch-Gordan coefficients. As a consequence, given a trajectory $\gamma : \mathbb{R} \to \mathbf{G}_{hom}$ and a matrix element $c : \mathbf{G}_{hom} \to \mathbb{C}$ among those given by (3.2) or (3.5), evaluating $c \circ \Gamma$ is not an unreasonable task for a computer − provided the ℓ_1 and ℓ_2 which appear in the formula for c are not too large. This was the starting point for the (alas unfruitful) analysis of experimental data related in the next chapter.

Let me close this section with a question on the full Galilei group $\mathbf{G}$. it possible to find a computer-friendly matrix coefficients for the unitary irreducible representations of $\mathbf{G}$ which are not in fact representations of $\mathbf{G}_{hom}$? I have not heard of any result in that direction.

Bibliography

[1] J.- M. Lévy-Leblond, *Galilei Group and Galilean Invariance*, in Group Theory and is applications (ed. E. Loebl), Academic Press (1971), pp. 221-299

[2] J.- M. Lévy-Leblond, *Nonrelativistic particles and wave equations*, Communications in Mathematical Physics **6(4)** (1967), pp. 286-311.

[3] E. Inönü and E. Wigner, *Representations of the Galilei group*, Il Nuovo Cimento **9(8)** (1952), pp. 705-718.

[4] V. S. Varadarajan, *Geometry of Quantum Theory*, second edition. Springer Science & Business Media, 2007.

[5] G. W. Mackey, *Imprimitivity for Representations of Locally Compact Groups I*, Proc Natl Acad Sci U S A. **35(9)** (1949), pp. 537–545.

[6] G. W. Mackey, *Theory of Unitary Group Representations*, Chicago Lectures in Mathematics, Chicago University Press, 1976.

[7] G. W. Mackey, *Unitary Group Representations in Physics, Probability, and Number Theory.* Benjamin/Cummings, 1978.

[8] J-M. Souriau, *Structure des systèmes dynamiques*, Dunod, 1970.

[9] J. Rawnsley, *Representations of a semidirect product by quantization*, Mathematical Proceedings of the Cambridge Philosophical Society **78(2)** (1975), pp. 345-350.

[10] N. Ja. Vilenkin, *Special Functions and Theory of Group Representations*, Izdat. Nauka, Moscow, 1965; Transl. Math. Monographs **22**, American Mathematical Society, 1968.

[11] B. Kursunoglu, *Modern Quantum Theory*, Freeman, San Francisco, 1962.

[12] W. Miller, *Lie theory and special functions*, Academic Press, 1968.

[13] G. Chirikjian and A. Kyatin, *An Operational Calculus for the Euclidean motion group with applications in Robotics*, Journal of Fourier Analysis and Applications **6(6)** (2000), pp. 583-605.

[14] R. Blattner, *Positive definite measures*, Proceedings of the American Mathematical Society **14(3)** (1963), pp. 423-428.

[15] B. Ørsted. Induced representations and a new proof of the imprimitivity theorem. Journal of Functional Analysis, 31(3), 355-359.

[16] N. Poulsen, *Regularity aspects of the theory of infinite-dimensional representations of Lie groups*, Ph. D. thesis, Massachussets Institute of Technology, 1970.

[17] E. Kaniuth and K. Taylor, *Induced representations of locally compact groups*. Cambridge Tracts in Mathematics **197**, Cambridge University Press (2013).

Chapter 6

Firing rate of Purkinje cells in the vestibulocerebellum and matrix elements of group representations

Contents

Abstract. This nonmathematical (and rather unusual) chapter reports on a collaboration with the Cerebellum group at the Institut de Biologie of the École Normale Supérieure. Matthieu Tihy, Guillaume Dugué and Clément Léna offered us a very nice opportunity to work on electrophysiological, single-unit recordings the activity of Purkinje cells in the nodulus of live, alert, and freely moving rats, with the hope that group theory, especially the matrix elements of unitary representations of the Galilei group described in Chapter 5, might be helpful in making sense of the data.

I used the results recalled in the last chapter to test for that hypothesis. Some of our findings are quite compatible with the existing literature on the cerebellum, and confirm that there are neurons which proceed to a linear combination of the semicircular canals' output. But alas! None of the ideas come from representation theory proved helpful in understanding the electrical activity of any neuron.

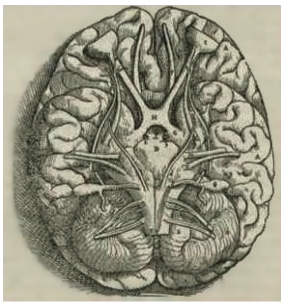
In this short chapter, I report on an attempt I made to study experimental recordings of the electrical activity of single neurons in the cerebellum of live rats. This is joint work with biologists Matthieu Tihy, Guillaume Dugué and Clément Lena, who are members of the Cerebellum group at the Institute for Biology at Ecole Normale Superieure, biostatistics expert Boris Gourévitch from the Neuroscience Institute at Paris-Saclay, and Daniel Bennequin.

I must say from the outset that the results I obtained are rather disappointing, and do not seem to indicate a particular relevance of the idea we put to the test− at least in the precise region in which we analyzed the neuronal activity. I nonetheless decided to include the report below to my thesis. My motivation for doing so, aside from the natural wish to say a few words about a year-long collaboration on real biological data (which taught me a lot), is that the idea−suitably adapted and more skillfully exploited−might prove useful in the study of other vestibular areas, and that the few nonnegative results I shall mention are rather nicely compatible with some prevailing ideas about the cerebellum, even though they ultimately necessitate no group theory.

This project would not have been possible but for the four years Matthieu spent working hard (night and day) on the experiment, the details of which are a genuine wonder to a mathematician. I am glad to thank him and the four other colleagues for the opportunity to work on real data. I would also like to thank Sophie Cachot for (among other things) getting me started on Python, and Sakina Madel for help with computers at IMJ-PRG.

1 The problem

The cerebellum is the part of the central nervous system that is located at the base of the skull, under the more famous cerebral hemispheres. About half of our neurons are located in it, though they are more tightly packed together there than they are in the cerebral cortex − the cerebellum's volume is about ten percent of the brain's total volume.



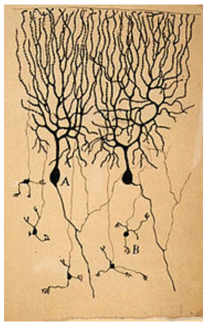
The base of the human brain, as drawn in 1543 by Vesalius. The cerebellum is at the bottom.

The cerebellum has long been known to participate in motor control, because damage to it usually results in clear perturbations of the motion-related behaviour: early and famous observations by Flourens and Dalton had made the idea well-accepted in the early and mid-nineteenth century, and Babinski had emphasized the cerebellum's contribution to equilibrium in 1899. Around 1904, Gordon Holmes confirmed their findings in a systematic examination of several wounded patients. And then many, many soldiers suffered cerebellar damage during the First World War; Holmes had many sad occasions to have evidence for the fact that the cerebellum has a key part to play in motion planning. His patients often lost their resistance to involuntary limb displacements or their ability to stand or walk, and frequently had much trouble coordinating their movements (*ataxia*) or performing any kind of rythmic or timing-dependent motor task.

Of the main cerebellar areas, the *vestibulocerebellum* is probably the most ancient. It receives inputs from the vestibular system and the visual system; in fact the path from the semicircular canals and the otoliths to the cerebellar neurons to be considered below, starting with the vestibulo-cochlear nerve and projecting to the vestibular nuclei in the brainstem and then directly to the cerebellum, involves two or three synapses at most. Damage to the vestibulocerebellum very clearly impairs equilibrium, and it also impairs eye motion planning – whether it be the vestibulo-ocular reflex or the ocular pursuit of a moving object.

Much is known on the vesitbulocerebellum, on its function and possible dysfunction, and much is being discovered as I write. Many mysteries remain, of course ; any information on the way the vestibulocerebellum processes the vestibular information should be of help in understanding how we perceive and plan our motions.

Some of the neurons in the cerebellum are very famous for their recognizable, and very well-documented, electrical behaviour: these are the *Purkinje cells*, big, neatly aligned cells with a large and elaborate dendritic arbor. Purkinje and Ramon y Cajal's nineteen-century studies (and drawings) of these cells have often been acclaimed as being among the foundations for all Neuroscience. The part these cells have to play in cerebellar function is crucial indeed: they are the only cerebellar neurons which afterwards project to the areas responsible for motor coordination.



A famous 1899 drawing of a Purkinje cell (pigeon) by Ramon y Cajal.

Several studies have provided insight into the way the electrical activity of Purkinje cells is modulated by the input from the semicircular canals and the otoliths. The fact that the canal and otolith signals *do* have an influence was observed in actual recordings of the electrical activity in 1976 [4].

Later studies (Fushiki 1997 [7], Sheliga 1999 [8]) focused on the influence of the canal-generated signals, and identified cells whose discharge is influenced by the *rotations* of the head, with a clear influence of the axis of rotation. It is only recently that the influence of rotations *and* translations together was examined (Yakusheva et al. 2007 [11], 2011 [12]). Several important points were made:

• While there *are* canal-only cells responding exclusively to the rotations (estimated to a third of the tested population), there are also *mixed* cells on which there is a clear influence of *both* the angular velocities and the linear accelerations of the head.
• Some of the neurons which process the vestibular information must succeed in converting the *self-centered* output of the end organs of the inner ear, which is insensitive to the position of the head relative to its surroundings, into a *world-centered* representation of the same signal, which is naturally essential for motor planning [1].
• In accordance with this, some cerebellar neurons (and others in the vestibular nuclei) participate in the global "computation" which must succeed in identifying the contribution of gravity within the otolith signal (Angelaki et al. 2004 [10], Yakusheva et al. 2007 [11]).

1. If one stimulates the vestibular receptors as subjects walk in the dark so as to evoke a given virtual rotation, as Fitzpatrick et al. did [9], the ensuing behaviour depends on the head's orientation at that time. This clearly indicates that the brain succeeds in keeping track of the head's orientation through time.

These studies focused on *passive* movements: a stereotyped movement was imposed to the animal's head and the effect on the electrical activity of the Purkinje cells was recorded. However, in the vestibular nuclei which receive projections from (and send projections to) the vestibulocerebellum, it has been very clearly established that neurons have sometimes dramatically different responses to voluntary (*active*) head movements than they have to constrained ones. Studies by Angelaki and Cullen's groups (and reviews by these authors) have emphasized this point; see [2, 5, 6]. It is tempting to ask whether such a difference can be observed in the cerebellum, too. This was the reason why Matthieu Tihy, Guillaume Dugué and Clément Léna recorded the Purkinje cells of the cerebellum during *active* navigation.

⋆⋆⋆

How are these current questions on the vestibular system and the cerebellum related with group theory and representation theory ? Why do I say that it is not absurd to imagine that representation theory might be helpful in discussing electrophysiological recordings of a wandering rat's cerebellum ?

The answer to these questions was given in [3]: we saw in the Introduction that the output of the canals and otoliths is a tangent vector to the homogeneous Galilei group $\mathbf{G}_{hom}$, and as a consequence, that there are natural ways to encode the head's motion by a smooth path in the homogeneous Galilei group − smooth maps

$$\begin{aligned} \mathbb{R} &\to \mathbf{G}_{hom} \\ t &\mapsto \gamma(t). \end{aligned}$$

If the electrical activity $t \mapsto \mathcal{A}(t)$ of a neuron depends on the output of the vestibular organs in a nontrivial, nonlinear way, and if there are neurons working in different reference frames and thus keeping track of the head's motion through time, it is a distinct possibility that there be a map

$$\mathfrak{F} : \mathbf{G}_{hom} \to \mathbb{R} \tag{1.1}$$

such that a significant part of the signal $\mathcal{A}$ can be explained by the composition $\mathfrak{F} \circ \gamma$. There is of course a mathematical difficulty in trying to find an explicit form for $\mathfrak{F}$, because it is a function defined on a six-dimensional manifold!

Now when G is a Lie group, noncommutative harmonic analysis provides a Fourier-like expansion of any function on G in terms of the matrix elements of the irreducible representations of G; I am now going to argue that if one has explicit, computer-friendly formulae for the matrix elements and if there is a function $\mathfrak{F}$ on G such that $\mathcal{A} = \mathfrak{F} \circ \gamma$, it is at least in principle possible (though in practice it can be difficult) to *find* $\mathfrak{F}$ from $\mathcal{A}$ and γ. In the case of neuronal activity, it makes it possible to look for the receptive profile $\mathfrak{F}$, even when the neuron lies deep into the cerebellum and $\mathfrak{F}$ is a function defined on a six-dimensional Lie group.

⋆

Suppose first that G is a *compact* Lie group. Then I recalled in the Introduction that if f is a continuous function on G and $\left\{\mathbf{c}^{\lambda}_{i,j}\right\}_{\substack{\lambda\in\widehat{G}\\ i,j=1..d(\lambda)}}$ is an exhaustive collection of matrix elements, then there is a square-summable family $\left\{\widehat{f}^{\lambda}_{i,j}\right\}_{\substack{\lambda\in\widehat{G}\\ i,j=1..d(\lambda)}}$ of complex numbers such that

$$f = \sum_{\substack{\lambda\in\widehat{G}\\ i,j=1..d(\lambda)}} \widehat{f}^{\lambda}_{i,j} \cdot \mathbf{c}^{\lambda}_{i,j}. \tag{1.2}$$

(the equality holds both in $\mathbf{L}^2(G)$ and pointwise). When f is smooth, the family $\left\{\widehat{f}_{i,j}^{\lambda}\right\}_{\substack{\lambda\in\widehat{G}\\ i,j=1..d(\lambda)}}$ usually decreases rapidly with the highest weight of λ, and a correct approximation in $\mathbf{L}^2(G)$ by a (finite) linear combination of matrix elements is likely to be found.
If one has computer-friendly formulae for some the $\mathbf{c}_{i,j}^{\lambda}$ and a recording of a path $\gamma : \mathbb{R} \to G$ and a map $\mathcal{A} \to \mathbb{R}$, then one can test whether there is an f such that $\mathcal{A} = f \circ \gamma$ by projecting f onto the finite-dimensional subspace of $\mathbf{L}^2(\mathbb{R})$ spanned by the chosen $\mathbf{c}_{i,j}^{\lambda}$s.

⋆

If G is not compact but is unimodular and type I, the unitary dual $\widehat{G}$ is no longer countable, but it is still true that there is an expansion of smooth and square-integrable functions on G in terms of matrix elements. Suppose a representative $(\mathcal{H}_\lambda, T_\lambda)$ is chosen for every class λ in $\widehat{G}$, a Hilbert basis (e_i^λ) is chosen for $\mathcal{H}_\lambda$, and write $\mathbf{c}_{i,j}^{\lambda}$ for the matrix element $g \mapsto \langle e_i^\lambda, T_\lambda(g)e_j^\lambda\rangle$. Then there is a positive measure μ on $\widehat{G}$, the Plancherel measure, with the property that every function f in $\mathbf{L}^2(G) \cap \mathbf{L}^1(G)$ can be decomposed as

$$f(g) = \int_{\widehat{G}} \left(\sum_{i,j} \widehat{f}_{i,j}^{\lambda} \cdot \mathbf{c}_{i,j}^{\lambda}(g) \right) d\mu(\lambda) \tag{1.3}$$

in which the family $\left\{\widehat{f}_{i,j}^{\lambda}\right\}_{\lambda,i,j}$ of complex numbers is such that the right integral has a meaning for every g.[2]

Although (1.3) does not say anything of non-square-integrable functions, it is also true that many of them be decomposed as combinations of matrix elements: to give a simple example, suppose M_f is a positive measure on $\widehat{G}$ with total mass one. Then as soon as $\left\{\widehat{f}_{i,j}^{\lambda}\right\}_{\lambda,i,j}$ is a collection of complex numbers such that the family $\left\{\widehat{f}_{i,j}^{\lambda}\right\}_{i,j}$ is square-summable for every λ (and thus can be empty save for a finite number of irreducible representations), then

$$f := g \mapsto \int_{\widehat{G}} \left(\sum_{i,j} \widehat{f}_{i,j}^{\lambda} \cdot \mathbf{c}_{i,j}^{\lambda}(g) \right) dM_f(\lambda) \tag{1.4}$$

is a smooth function on G, but it is neither integrable nor square-integrable in general. When G is a semidirect product $H \ltimes A$ with H compact and A noncompact abelian, functions f on G which are obtained by extending a function φ on H to G (setting $f(h,a) = \varphi(h)$) satisfy (1.4) with a measure M_f concentrated on the subset of $\widehat{G}$ gathering the representations of H; in that case the expansion (1.4) reduces to the Peter-Weyl expansion (1.2) for smooth functions on H.

⋆

Because the equivalence classes of irreducible representations can be viewed as "generalized frequencies" in view of the above formula, and because we saw that neurons in the auditory system and the visual system each work "one frequency at a time" for signal processing, it is very appealing to speculate that if a "receptive profile" $\mathfrak{F}$ as above exists (see (1.1)), it might actually be a linear combination of matrix elements of a *single* irreducible representation. A closer look at the concrete formulae (10) in VI.3.2 and their physical interpretation makes this suggestion reasonable in the case of the Galilei group: the continuous parameter R which is necessary to identify an irreducible representation appears

2. To be precise, when λ is fixed the operator $\widehat{\mathbf{f}}_\lambda$ whose "matrix" in the chosen basis (e_i^λ) is $(\widehat{f}_{i,j}^{\lambda})_{i,j}$ is Hilbert-Schmidt, and the map ($\lambda \mapsto$ Hilbert-Schmidt norm of $\widehat{\mathbf{f}}_\lambda$) is integrable with respect to μ.

there only as a scale factor for the dependence of the matrix elements on the velocities' norm, and fixes the range of velocities to which the neuron is likely to be sensitive.

It might not be out of place here to recall that in contrast to the fact that the representations of the rotation subgroup have Plancherel measure zero in $\widehat{\mathbf{G}_{hom}}$, the population of vestibular neurons working solely with the rotations is estimated to about 30% of the total. So in spite of its lesser mathematical generality, (1.2) is well worth testing − if only because it is very much more easily tested on a computer.

This chapter is a report of my attempt to study an experimental recording of Purkinje cells in the vestibulocerebellum of live rats with the help of (1.2) and (1.4). I tested whether some neurons can be said to have a "receptive profile" $\mathfrak{F}$ (see (1.1)) given by the version of (1.4) in which M_f is either concentrated on the subset of $\widehat{\mathbf{G}_{hom}}$ gathering the representations of the rotation subgroup, or else is a Dirac mass concentrated on a single infinite-dimensional representation with helicity zero.

2 The experiment

When I started working on this project, the experimental data had for the most part already been obtained by Matthieu Tihy and Guillaume Dugué. Here is the challenge they had faced and met.

Consider a single Purkinje cell in the vestibular cerebellum of a rat
who is moving freely, in daylight or in darkness.

Keep track of each moment when the cell emits a spike.
Keep track of the position, angular velocity, and linear acceleration of the head
*at that **precise** moment.*

Matthieu Tihy and Guillaume Dugué had worked with a group of twelve rats aged three to four months. The rats had first received implantatory surgery [3]:

• a tetrode meant to record the electrical activity of single neurons was implanted so as to have its tip in the appropriate region of the cerebellum−the tetrode was devised and built at the Institute, and the region is "lobulus X" (the nodulus), a ventral region in the cerebellum just above ventricle V4. The tetrode was lodged within an Institute-built base (fixed atop the head) from which the position of the tetrode's tip within the cerebellum could be adjusted. The tip's position could be very finely adjusted until a pattern of electrical activity typical of a single Purjinke cell appeared (thus guaranteeing that a single neuron was being recorded).

• a platform was fixed upon their head; the platform could host an inertial sensor (this was a numerical microelectromecanic sensor devised and build within the lab) meant to record the angular velocities and linear accelerations of the head during motion. The sensor conveyed the information to a (National Instruments) software for which an interface had been specifically devised.

3. The surgery was of course careful and mindful of international ethics standard well-known to biologists. Matthieu Tihy's thesis contains a description of the surgical procedure.

Each rat was free to move at will in an open space a meter above the ground. During this time,

• the tetrode kept track of each moment when the cell emitted a spike;

• the inertial sensor recorded the angular velocity and linear acceleration of the head every five milliseconds.

Very close attention was paid to the synchronization between both recording systems; it was crucial that the synchronization be perfect, because one of the key points to analyze was the time delay between the current state of motion and each cell's reaction to it. As I recalled, it is well-known that some neurons in the recorded area have a key role in motion *planning*, and thus might be found to have "anticipatory" responses. This was a crucial point to discuss and a source of very hard challenge for the Institute's team; unforeseen challenges kept coming up to the very end.

Matthieu Tihy then performed the necessary, but often delicate, integrations. He obtained

(a) An estimation of *the average electrical activity of the neuron at each time t* (this estimation was obtained from the spikes' firing dates by moving averages).

(b) A value for *the orientation of the head at each time t*, relative to its orientation at the start of the recording session; the successive orientations were represented by successive rotations sending the initial orientation to the current head orientation, with each rotation encoded by a unit quaternion (see below). For this, Matthieu used a recent version of an algorithm of the kind usually implemented in aircrafts. The accuracy of the numerical integration (which is rather delicate to perform because the head orientation lives on a compact group rather than an affine space) was controlled in a very simple way: the rats were filmed as they moved, and the algorithm-predicted orientation of the nose at time t could then be compared with the actual orientation at time t.

(c) A value at each time t for the *close-to-the-identity rotation* sending the position of the head at t to the position of the head shortly after t (see below). This was again encoded by a unit quaternion.

(d) An estimation of the linear acceleration at time t (this came directly out of the inertial sensors)

(e) An estimation of the linear acceleration *with the contribution of gravity substracted* (recall that the sensors are tied to the head; the estimation then needed the estimation of the head's orientation at each time, as performed in (a)).

(f) An estimation of the linear velocity at time t (after an ordinary numerical integration).

I will refer for brevity to "the orientation and velocity in an allocentric frame" to mean the path $\gamma_{\text{ext}} : \mathbb{R} \to \mathbf{G}_{hom}$ defined by (b) and (f). In the reference frame attached to the rat's head, the head does of course not move; it is nevertheless useful to imagine that neurons working in a self-centered frame can work with the "small" motions undergone by the head (thus using both the reference frame attached to the head at time t and that at time $t + \delta$, with δ a short time delay. If $\mathfrak{D}_t$ is the element of the Lie algebra of $\mathbf{G}_{hom}$ defined by (c) and (d) or (e), I will refer to "the small motions in an egocentric frame" to denote the path $\gamma_{\text{int}} = t \mapsto \exp_{\mathbf{G}_{hom}}(\delta \mathfrak{D}_t)$, where δ is a short time delay (we used 20ms, which seemed both biologically and numerically reasonable).

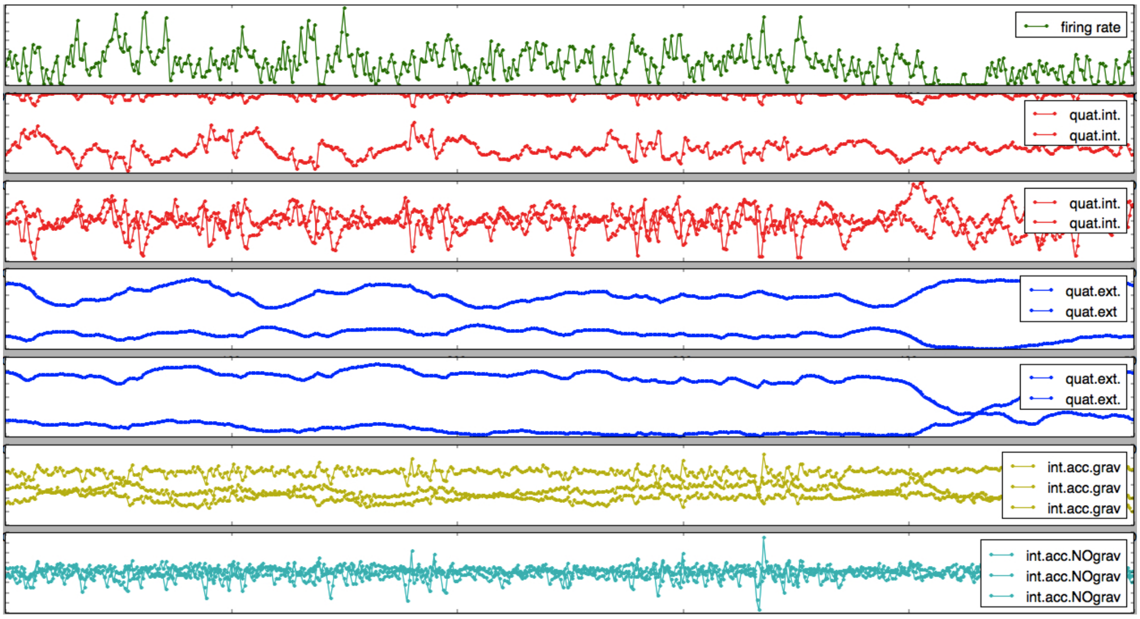


Figure 1: This is a ten-second-long extract of the raw data for one cell (datafile #2, between 20' and 30'). The red (resp. blue) curves display the four components of the unit quaternion encoding the orientation of the head in an egocentric (resp. allocentric) frame, and the yellow (resp. cyan) curves display the three components of the linear acceleration of the head at time t (resp. of the linear acceleration with the contribution of gravity substracted).

For each rat, recording sessions took place daily over a period of two weeks. Some of these sessions took place with the light on, some others took place in the dark. Remembering that many existing studies on the cerebellum used passive movements in which a stereotyped trajectory was imposed to the head, some special recording sessions took place: in these the animal was not moving freely, but the rat and its head were forced to follow a predetermined course.

Matthieu Tihy extracted from this a set of 90 datafiles containing the estimations (a) to (f) every five milliseconds over a period of about five minutes, and he sent the files to me. He knew which rat, which recording session and which cell were associated to a given file, but I did not. One of our secondary aims was to see whether our numerical tests made it possible to identify which files were attached to a given cell (and whether there was a difference between passive and active movements).

3 The numerics

Given the discussion in Section 1 and the available data, there were several very natural questions we could turn to:

• Are there neurons whose electrical activity can be partially understood as a function of γ_{ext} or γ_{int} with the help of (1.4) ?
• If so, how anticipatory is their response ? What is the time delay for their reaction to the sensory input [4] ?
• Can we identify "mixed" cells in the recorded population, and do they use matrix elements

4. The time delay might be negative if the neurons actually anticipate on the trajectory.

of infinite-dimensional representations of $\mathbf{G}_{hom}$? If so, is it the otolith signal that is used, or is it the gravity-free version of the otolith signal ?

⋆

Here is a short slogan for the tests I performed :

For each of the 90 files containing descriptions of a rat's head motion and a neuron's electrical activity, view the rat's head motion as a trajectory in the homogeneous Galilei group $\mathbf{G}_{hom}$, evaluate the matrix elements of section VI. 3.2 along that trajectory, and try to see whether the electrical activity can be understood with their help.

To describe the strategy a bit more precisely, let me start by recalling that a unit quaternion $q_0 + q_1 I + q_2 J + q_3 K$ can be viewed as the unitary matrix

$$\begin{pmatrix} q_0 + iq_1 & q_2 + iq_3 \\ -q_2 + iq_3 & q_0 - iq_1 \end{pmatrix}.$$

I wrote a Python code for evaluating the matrix elements in VI.3.2. Because the matrix elements of $SU(2)$ detailed there become homogeneous polynomials in the quaternion's four coordinates, it was rather easy to evaluate the matrix elements of representations actually extended from representations of the rotation group (one needs only specify the triple (ℓ, m, n) corresponding to the desired coefficient).

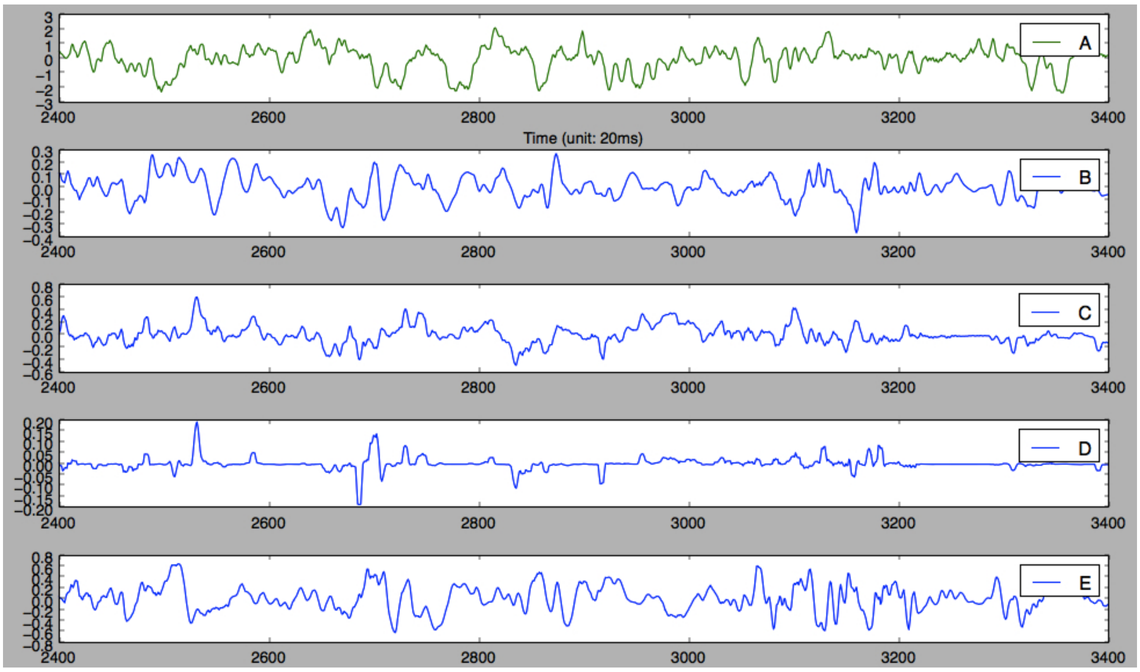


Figure 2: This is a display of some matrix elements extended from matrix elements of $SU(2)$, evaluated over twenty seconds (datafile #28, 48' to 68') over the orientations "in an egocentric frame" (small motions of the head). Curve **A** is the gaussianized activity of the cell (see below). Curves **B** and **C** are matrix elements of the irreducible representation with $\ell = 1/2$, and thus are just two components of the unit quaternion. Curves **D** and **E** are coefficients matrix elements of the irreducible representations with $\ell = 1$ and $\ell = 2$, respectively. What is displayed is a low-pass filtered version in which the frequencies above 10Hz were cut off (see below).

When I worked with the "small motions in an egocentric frame", that is, with the path $\gamma_{\text{int}} : \mathbb{R} \to \mathbf{G}_{hom}$ taking values only in a small neighbourhood of the identity, I took account of the fact that the path can be lifted to the universal covering of $\mathbf{G}_{hom}$, and tested for the matrix elements of *all* irreducible representations of $SU(2)$, not just those which factor through the covering morphism $SU(2) \to SO(3)$. This means that I allowed for matrix elements which were polynomials of *odd* order in the quaternions' four components; in particular, I left open the possibility that the neuron's response be a *linear* combination of the quaternion's four components. We shall see that this does happen, and that among biologists it has always been a very well-acknowledged possibility that it should.

⋆

The calculation of the special functions corresponding to infinite-dimensional representations was rather more laborious, because a continuous parameter R and an integer ("helicity") s had to be specified in order to single out one irreducible representation, and four integers $(\ell_1, \ell_2, m_1, m_2)$ had to be specified in order to point to one matrix element among those in VI.3.2. Because the calculation had to terminate in a reasonable time, I had to keep to values of ℓ_1 and ℓ_2 below 2; I also had to keep with $s = 0$ – in any event it seemed more likely that in the event that a neuron worked with these arcane special functions, the representations of helicity zero were the most natural candidates. With these choices, one gets 81 special functions for each value of R.

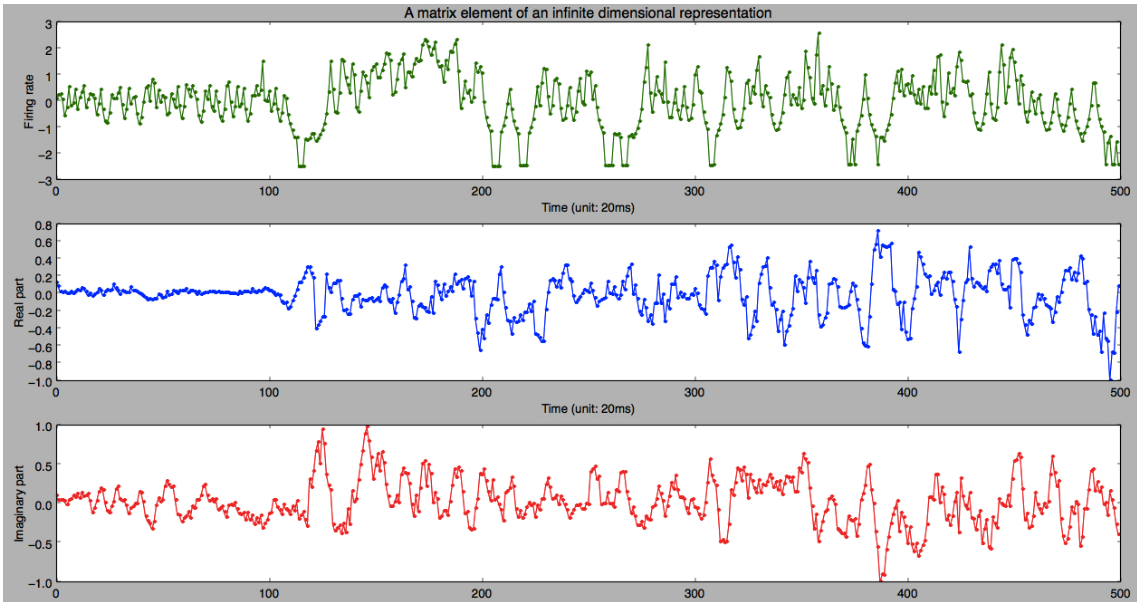


Figure 3: This is a display of a matrix element of the irreducible representation of $\mathbf{G}_{hom}$ with continuous parameter R=6000. and helicity zero, evaluated over twenty seconds (datafile #28, 48' to 68') on the orientations "in an egocentric frame" (small motions of the head). The matrix element is the one with $(\ell_1, \ell_2, m_1, m_2) = (1, 1, 0, 1)$. The blue curve is the real part, the red curve is the imaginary part, and the green curve is a display of the cell's firing rate (gaussianized and smoothed version, see below) at that time.

What the test is.

Given the formulae in VI.3.2 and the appearance there of global factors involving the value $j_k(Rv)$ of a Bessel function, the continuous parameter R specifies the range of velocities or accelerations above which the matrix elements become very close to zero. I chose this parameter by looking at the raw data and the output of VI.3.2, while keeping track of the physical plausibility of the result (the chosen R should keep the accessible range of velocities daily-life-like). Given an infinite time to run the program, it would have been possible to optimize on R through an a automatic procedure, but I had to strive for brevity (and as we shall see, it turned out that no value of R came close to explaining the data).

Choose a set $\{F_j\}_{j=1...N}$ of functions built from group-theoretical matrix elements.
For a given time increment δ,
calculate the $\mathbf{L}^2$-distance between $t \mapsto \mathcal{A}(t)$ and Span$\{t \mapsto F_j(t - k\delta)\ ,\ j = 1..N\}$, $k \in \mathbb{Z}$.
Find the time delay $k\delta$ for which it is minimal.
Look at the projection of $\mathcal{A}$ on the corresponding vector space. Does it look like $\mathcal{A}$?

Of course calculating the projection and the distance can be done through the least-square algorithm, and that part is naturally implemented in Python. This makes the test rather reasonable in computational terms; as we shall see, however, using the $\mathbf{L}^2$ distance is very far from being ideally suited to the nature of the data we explored.

Convolution with an appropriate time kernel.

It is customary in Neuroscience to assume that when a neuron responds to the information given by a sensor around time τ, its response can depend on the information at neighbouring instants in a variety of ways. A standard way to take this remark into account is to introduce a *time kernel* $s \mapsto \kappa(s)$ which is centered at zero and rapidly decreasing, and to test the correlation between $\mathcal{A}$ and the convolutions $\tau \mapsto \int F_j(s)\kappa(\tau - s)$ rather than that between $\mathcal{A}$ and the F_js. Among the usual candidates for κ are

– Gaussian functions centered at zero; using such a kernel is not unlike testing the hypothesis that the neuron uses a low-pass-filtered version of the F_js; the Gaussian kernel's width yields an indication of the frequency threshold above which the neuron can no longer take the information into account.

– Derivatives of Gaussians centered at zero; this is an appealing possibility, because these functions are either negative at t for $t < 0$ and positive at t for $t > 0$, or positive at t for $t < 0$ and negativeat t for $t > 0$. In that case the "variable" F_j has an inhibitory-then-excitatory effect on the neuron's activity.

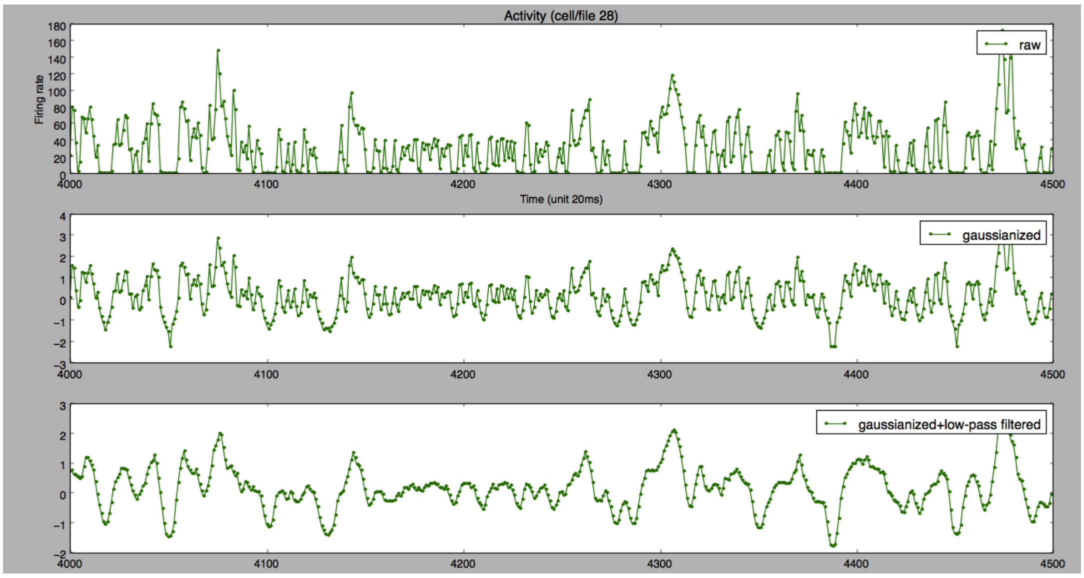
Of course it is necessary to use the same kernel for every j, otherwise the risk of "overfit" is very high. I tested five kernels: a Dirac mass, a Gaussian function with standard deviation corresponding 20ms, a Gaussian function with standard deviation corresponding to 80ms, and the derivatives of these two Gaussian functions.

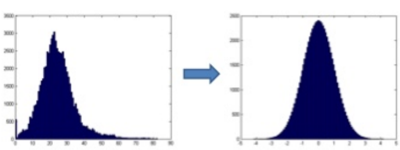
Firing rate gaussianization and filtering.

Another custom in Neuroscience is to remark that when one extracts a firing rate from the spikes' emission dates and when the firing rate is then considered as a random variable, its distribution depends on delicate physico-chimical phenomena and might not be relevant for understanding the correlation between the response and a real-valued function with another distribution. In our case the mean value of the firing rate is obviously positive and its distribution is Poisson-like, but the distribution of values for the angular velocities is approximately Gaussian, and that of accelerations is approximately Gaussian as well (at least when gravity is substracted).

In addition, one should keep open the possibility that the behaviour of $\mathscr{A}$ might depend on the tested variables F_j only in a band of frequency. Very high frequencies in the neuron's electrical activity are rather unlikely to be part of the neuron's response to head movements whose frequency cannot exceed a reasonable upper bound. In particular, it is not absurd to suppress the very-high or very-low frequency components of $\mathscr{A}$ before testing the $\mathbf{L}^2$-distance.

I thus followed a rather frequent procedure in composing $\mathscr{A}$ with an increasing diffeomorphism of $\mathbb{R}^+$ which turns the distribution of the firing rate into a standard Gaussian one, and then filtering the result so as to keep only the frequency components between 0.5Hz and 10Hz.





The first curve above is a plot of one cell's firing rate over ten seconds; the second is the gaussianized version, and the third is what the gaussianized version looks like after one cuts off the frequencies above 10Hz.
On the left is a picture of that cell's firing rate distribution; the Gaussianization step consists in composing the firing rate's values with the diffeomorphism of $\mathbb{R}^+$ *sending the left curve to the right curve.*

Of course I tested several frequency cut-offs and checked that while my algorithms' results depend on the chosen frequency range, the precise value for the cut-off frequencies is not so crucial as to have the results change drastically if one of the cut-off frequencies is changed marginally.

4 The results

Let me recall that γ_{ext} is the notation for the map $\mathbb{R} \to \mathbf{G}_{hom}$ recording the orientation and velocity in an allocentric frame, and γ_{int} is the notation for the map $\mathbb{R} \to \mathbf{G}_{hom}$ recording the "small motions of the head" in an egocentric frame. A look at the firing rate with the naked eye, and a look at the values of the matrix elements along γ_{ext}, reveals that the matrix elements unsuprisingly have much slower variations than the firing rate has − recall that the components of γ_{ext} record the (visible!) motion of the animal, so that in spite of the small-amplitude tremor characteristic of a rat's head motion, they usually vary much less quickly than the firing rate does (compare the blue and green curves on Figure 1).

If we are to make an attempt at explaining the activity of a cell with the help of matrix elements of $\mathbf{G}_{hom}$, the not-very-low frequency components should be related to the orientation and velocity in an egocentric frame rather than an allocentric one, thus to the matrix elements $F_j \circ \gamma_{\text{int}}$ or else to the *derivative* of the matrix elements $F_j \circ \gamma_{\text{ext}}$. As I said above, applying a derivative-of-gaussian time kernel to $F_j \circ \gamma_{\text{ext}}$ is a way of testing the latter possibility.

I thus separated the low-frequency components and the higher (but not above 10Hz) frequency components of the activity, then proceeded to two separate comparisons:

• I tried to fit the gaussianized activity, low-pass *and* high-pass filtered so as to keep only the frequency components between 1Hz and 10Hz, , and the linear combinations of $F_j \circ \gamma_{\text{int}}$ (various time kernels) or $F_j \circ \gamma_{\text{ext}}$ (derivative-of-gaussian time kernels with *small* time amplitude) for various sets of F_js.

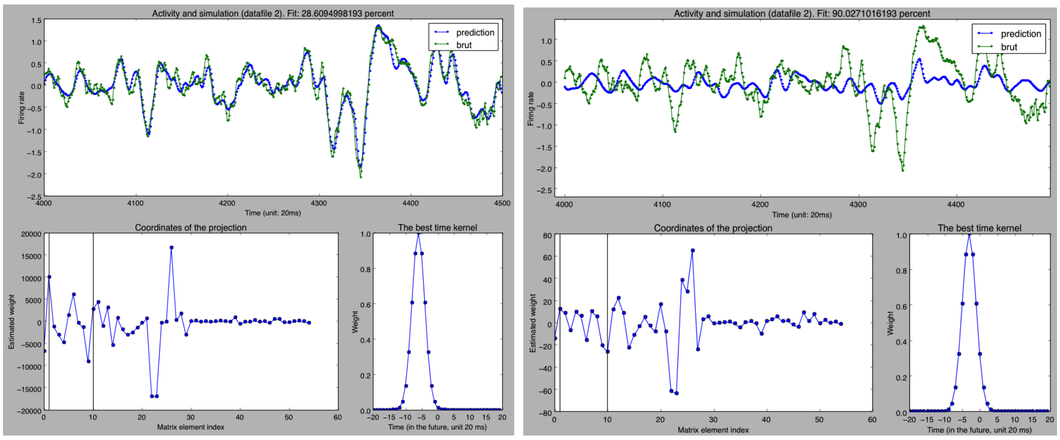
• I tried to fit the gaussianized activity, *low*-pass filtered so as to keep only the slow variations (below 1Hz), with the linear combinations of $F_j \circ \gamma_{\text{ext}}$ (gaussian or derivative-of-gaussians with *large* time amplitudes.

High-frequency components: testing with rotations

Here the set of chosen matrix elements F_js will correspond to the irreducible representations of $SU(2)$ whose highest weight parameter ℓ is between 0 and 2 (as we shall see, going further up in the highest weight order does not improve the results significantly): since the dimensions of the corresponding vector spaces are 1, 2, 3, 4, 5, this means that I used 1+4+9+16+25=55 special functions.

Let me write $s\delta$ for the time shift with delay δ: $\mathcal{A} \circ s_\delta$ is $t \mapsto \mathcal{A}(t-\delta)$, and $\tilde{F}_j$ for the convolution $F_j \star \mathcal{K}$, where $\mathcal{K}$ is one of the time kernels discussed above.

I wrote a Python code for choosing a δ-range, calculating the $\mathbf{L}^2$-distance between $\mathcal{A} \circ s_\delta$ and $\tilde{F}_j$ for every δ in the chosen range, finding the time delay that minimizes the distance, and displaying the $\mathbf{L}^2$-projection together with the 55 coefficients giving the coordinates of the projection in $\tilde{F}_j$-space. Here is what the output looked like in two closely related cases; beware that the superb coincidence between the two curves on the left is a fallacy, for reasons to be discussed immediately.



Although the left prediction looks excellent, it does in fact a very poor job: the only difference between the left curve and the right curve is that the left curve was obtained by *running the test over an ten-second extract* of the datafile, while the right curve *displays* a ten-second extract of the output of the same test, when *run over a two-minute extract*. What happened on the left is known as "overfit": there are enough functions to fit almost anything reasonable, but the excellence of the fit does not point to any biological phenomenon.

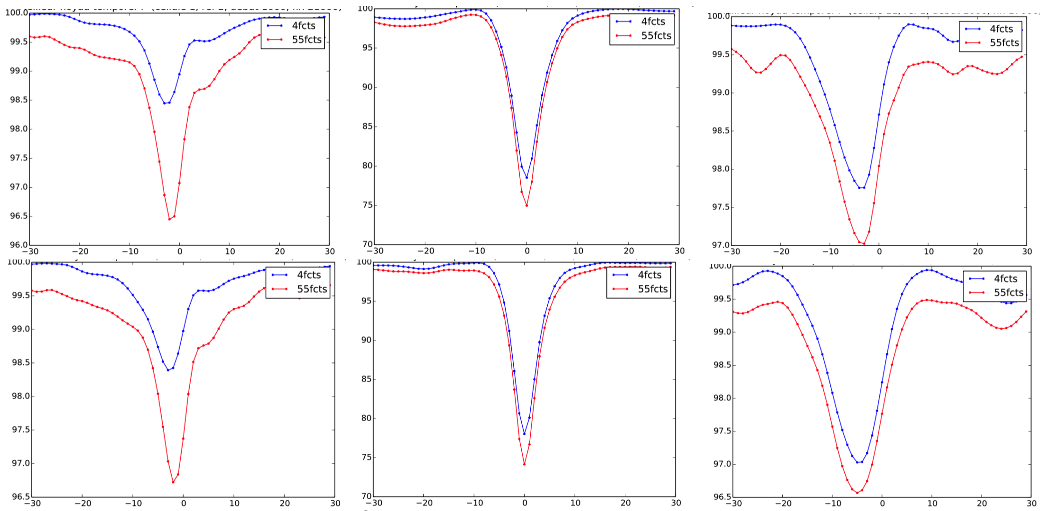
A general conclusion, which has a clear influence on the reliability of the resuts, is that the tests should be run over long enough extracts of the data: when the raw data looked clean enough for the test to be run over the whole datafile, I did so.

⋆

With this caveat, I can now describe some of the results.

As regards time kernels, I found that Gaussian kernels sometimes yield nice results, and the optimum standard deviation of the gaussian seems to be around 40ms (thus cutting off the frequencies above 10Hz). Derivative-of-gaussian kernels, however, lead to much poorer results. In the rest of this subsection, $\mathcal{K}$ will thus be a Gaussian function with standard deviation 40ms.

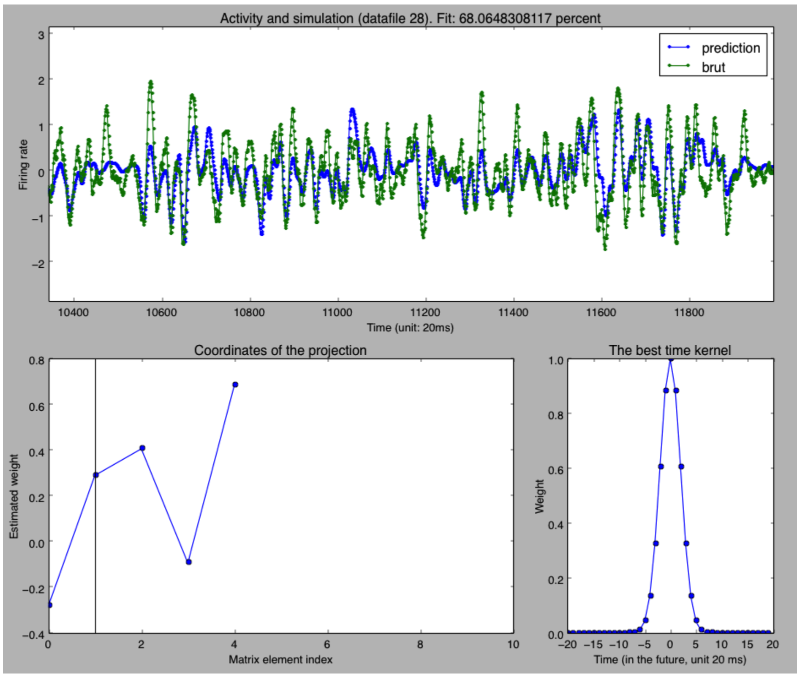
It was rather striking to see how the $\mathbf{L}^2$-distance and projection varied with δ: a first, reassuring fact was that it was usually maximal for large δ (whether positive or negative), and decreased as δ approached an optimal value. When the sample was long enough, the projection was typically zero for large δ, but sometimes the $\mathbf{L}^2$-distance decreased rather spectacularly and not-unrealistic-looking predictions turned up as δ approached the optimal value. A more interesting second fact was that both the shape of the curve and the location of the optimum varied as the datafile changed, but some patterns were common to several datafiles. *In fact, the datafile whose curves had the same shape were those attached to a single neuron.* The optimum time delays were all realistic, ranging from 0 to 100-120 ms in the past. This is a strong indication that the recorded neurons actually work with the canal-sensed input, though not necessarily in a direct manner, and this also gives an indication of the time that elapses before each neuron reacts to the canal-sensed signal.



These are plots of the relative $\mathbf{L}^2$ *distance between the delayed activity and the space spanned by four (blue curves) and 54 (red curves) matrix elements of* $SU(2)$*, viewed as a function of the time delay. Each test was run over the whole file, which comprised five to seven minutes of data. The files correspond to three (and only three) distinct cells, and the correspondence is easily imagined. I shall comment on the high relative* $\mathbf{L}^2$ *distances and the small difference between the blue and red curves soon.*

The way the $\mathbf{L}^2$-distance between $\mathcal{A} \circ s_\delta$ and the subspace generated by the $\tilde{F}_j \circ \gamma_{\text{int}}$
varies with the time delay δ
depends only on the neuron (not on the chosen extract):
the neuron can be identified from it.

The time-delay-dependence profiles clearly single out one neuron for which the decrease of the $\mathbf{L}^2$-error (as one approaches the optimal time delay) is spectacular in speed and in depth. A look at the $\mathbf{L}^2$ projection that comes out of the test reveals that for each of the datafiles corresponding to this neuron, the prediction is qualitatively very close to the actual activity over the whole datafile.



This is a thirty-second long ***extract*** *of the output of the test,* ***run over five successive minutes****. The blue curve is a prediction of the activity through a linear combination of the four coordinates of the quaternion in* γ_{int} *(and a constant due to the fact that one of the coordinates keeps close to one rather than zero). Since the prediction might not look perfect to a mathematician, I should emphasize that only four functions are used here, that the raw data were extremely delicate to obtain, and that the electrical activity of a neuron is a subtle thing: having a prediction for which the extrema are approximately in the right place, and which is valid over a long period of time with few functions, is much more important than having the value of the extrema right.*

Even more interestingly, the coefficients that come out of the test were the *same* for each of the four datafiles corresponding to the given cell. For this neuron, the prediction can thus be considered to reproduce the activity over a period of about *twenty minutes.* Since the quaternion's coordinates, when the quaternion is very close to the identity as it is here, are approximately affine in the Euler angles of the rotation, a linear combination of the quaternion's coordinates is a linear combination of the angular velocities sensed by the semicircular canals and sent to the cerebellum *via* the vestibulo-cochlear nerve. Thus the result for this neuron is very positive:

For **one** of the cells, the activity was very strikingly well-reproduced using a linear combination of the quaternion's four entries.
This cell's activity closely follows a linear combination of the semicircular canals' output.

This is very nicely compatible with the existing literature on Purkinje cells. The time delay is very short, in fact short enough that the neuron can be considered to participate in the planning of motion rather than its recording.

I will hasten to say that the interpretation of the activity as a linear combination of the angular velocities, while biologically nice, means that this particular instance of my group-theory-based test is one in which group theory is actually irrelevant! We set out looking for nonlinear receptive profiles, and identified one neuron for which there is a perfectly linear receptive profile.

In fact, on the above figures showing the decrease of the $\mathbf{L}^2$ error as one approaches the optimal time delay, the lack of a significant difference between the blue and red curves is witness to the fact that when there is a function of the rotation part of γ_{int} which can help describe the activity, there is very little to be gained by allowing it to be nonlinear.

From this point to the end of the chapter, I shall briefly describe a sequence of quite negative results which make it clear that the work undertaken in this chapter is an overall failure, one which would have justified passing over the attempt in silence. We did not see anything new with the help of group theory; it would be dishonest not to say so.

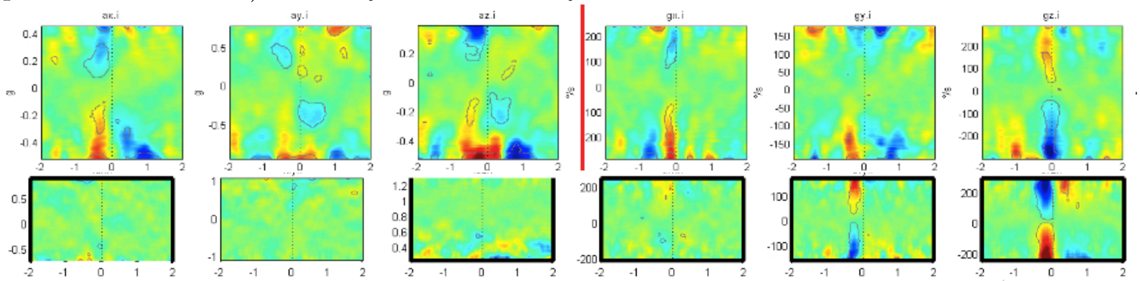
High-frequency components: testing for a mixed behaviour with infinite-dimensional representations

If we keep to linear combinations of the quaternions describing the rotation part of γ_{int}, although there *is* a decrease in the $\mathbf{L}^2$ error as one approaches the optimal time delay, the decrease is very far from that seen on the above neuron, and the projection is far from looking like the activity.

Now, our whole investigation was motivated by the possible existence of some cells blending the canal and otolith signals so as to work with both rotations and translations at the same time. The matrix elements of infinite-dimensional representations blend rotations and translations together, and as we saw on Figure 3, their value along γ_{int} is not absurdly unlike the electrical activity of some cells.

So are there some mixed cells in the recorded population, and if so, can their activity be understood with the help of matrix elements of infinite-dimensional representations ?

The answer to the first question is that yes, some of the recorded cells are likely to work with both rotations and translations. Witness to this is an analysis conducted by Boris Gourévitch, who drew the *spike-triggered receptive fields* of the cells. If $\mathcal{X}$ is an arbitrary function of time and δ is a real number, one can sort the values of $\mathcal{A}(t+\delta)$ according to the value of $\mathcal{X}$ at t; the map from the product $\{$values of $\mathcal{X}\} \times \{$ values of $\delta\}$ to $\mathbb{R}$ thus produced is called a spike-triggered receptive field. Below are plots of such maps for two distinct cells when δ, on the horizontal axis, ranges from -2 seconds to 2 seconds, and when $\mathcal{X}$ is one of three components of the linear acceleration or the angular velocity in or around a given direction, expressed in an egocentric frame. A red (resp. blue) point at (x,δ) signals that the cell's firing rate at t is systematically higher (resp.lower) than average when $\mathcal{X}$ takes the value x at $t+\delta$. One of the cells clearly looks influenced by both the linear accelerations and the angular velocities, while the other (discussed in the previous subsection) is clearly a rotation-only cell.



In the upper row are spike-triggered receptive fields for the three components of the linear acceleration in an egocentric frame (three left images) and the angular velocity (three right images) in an egocentric frame; in the lower row are the receptive fields for the one cell described in the previous paragraph.

⋆

The answer to the second question is no. I tried to fit the activity with a linear combination of the matrix elements of a single infinite-dimensional representation with helicity zero, whose continuous parameter I tried to adjust by hand (though it proved difficult, for there was no obvious choice). In many cases, a few of the individual matrix elements had activity-like variations (see Figure 3), but none of the linear combinations of the signals I obtained came close to the activity of the cell over a long enough period of time. For almost every cell it was possible to obtain a good prediction over various periods of perhaps forty seconds, but taking different samples of the same cell's activity and comparing the coordinates of the projection revealed a clear case of "overfit": no linear combination of matrix elements was tailored to one cell. This happened with both versions of the linear accelerations in an egocentric frame – the one with gravity substracted and the "raw" otolithic output. This happened for each of the time kernels I tested.

Rather than discussing the failed attempts we made at seeing something in the data, it is perhaps best not to make this chapter too long and acknowledge the failure:

Alas!
None of the matrix elements of infinite-dimensional representations of $\mathbf{G}_{hom}$
proved to be of any help in discussing the electrical activity of the neurons
even when they had been identified as combining rotations and translations together.

Low-frequency components

With the high-frequency components out of the picture, the next thing to look for was whether the slower variations of the activity could be explained with the help of matrix elements evaluated on γ_{ext}.

Rotations only. A first, reassuring remark is that the linear combinations of the quaternion part of γ_{ext} are very far from being able to reproduce anything like the slow components of the activity of any cell, even over short periods of time.

However, the analysis of spike-triggered receptive fields had shown that some of the cells' slow variations of the activity had nontrivial correlations with the rotation part of γ_{ext}, suggesting that these cells do work in an allocentric frame. So, can group theory help us find receptive fields ?

Turning to polynomials of higher degree, it might have been nice to see that for some cells we *could* obtain a projection which was, over every sample lasting one to two minutes, not unlike the activity. Does it correspond to the idea that the neurons somehow use γ_{ext}, and do so in a nonlinear way ? We did have hopes, but the few positive results turned out to have serious enough drawbacks to discourage any positive conclusion:

– The coefficients that came out depended on the chosen sample: no prediction stood its ground for significantly more than two minutes;

– When I converted the rotation part of γ_{ext} to three Euler angles and tried to compare the results with the projections on the subspace spanned by the polynomials of degree ≤ 3 in the Euler angles−94 special functions which have absolutely no intrinstic meaning and depend on an choice of axes uncorrelated to the canals' planes− the prediction was virtually indistinguishable from that obtained with the matrix elements.

Alas for group theory and for the idea that the neurons might use it to work with γ_{ext}, all that came out of this test thus seems to be due to the Stone-Weierstrass theorem on polynomial approximation.

Infinite-dimensional representations. Once again, I tried to use the matrix elements of infinite-dimensional representations of $\mathbf{G}_{hom}$, evaluated along γ_{ext} and with a continuous parameter adjusted by hand; once again, that was to no avail. Unlike what had happened for the higher frequency components, here the analysis of spike-triggered receptive fields had already revealed that it was unlikely that any one among the recorded cells used an allocentric coding of *both* rotations and translations. So this new failure is biologically not surprising, and it is perhaps not useful to dwell on it.

Alas!
When the head's orientation and velocity "in an allocentric frame" were used, although there were indications that some cells worked in these coordinates, there does not seem to have been any cell whose activity could be understood with the help of matrix elements of $\mathbf{G}_{hom}$.

5 Three concluding remarks

Looking back on the attempt, several weaknesses of our approach spring to mind. I would like to end this chapter by dwelling on three of them; the first is mathematical, the second is computational, and the third is biological.

• The starting point for this study was the fact that there are quite concrete formulae for the matrix elements of $\mathbf{G}_{hom}$. This point seems worth emphasizing, because matrix elements of unitary representations are well-known as a theoretical provider for special functions, and because it is clear that those of $\mathbf{G}_{hom}$ are likely to be of use in building or understanding motion-related systems. However, the fact that each individual matrix element in VI.3.2 is rather computer-friendly does not mean that there are no technical difficulties in dealing with many matrix elements at the same time. When there is no single irreducible representation to privilege, or when the one representation to privilege is not known, making no arbitrary choice can become quite challenging (here I *did* make arbitrary choices in despair).

• I used the least-squares algorithm to fit the data with linear combinations of matrix elements, because as a nonexpert it seemed the obvious thing to use when one has linear combinations of special functions to fit the data with, and because the $\mathbf{L}^2$-distance is frequently used in the analysis of neural activity (it is called usually the "R^2 parameter"). But after seeing what the predicted curves were when the predictions were not completely absurd, it is not obvious at all that the $\mathbf{L}^2$-distance is the best distance to use here: when the curves oscillate constantly for twenty minutes, changing one of the curves through a very slight deformation of the t-axis (while leaving the other curve unchanged) can lead to a dramatic increase in the $\mathbf{L}^2$ distance, whereas the prediction will qualitatively still account for the behaviour of the other curve in a way satisfactory for Neuroscience. Is there a universally-acknowledged better distance to use, one which it would have been computationally reasonable to test here ?

• The region in which the recordings took place is a very important one for the control of motion, and it is especially so because it is *close* (in terms of the number of synapses necessary to reach it) to the canals' and otoliths' output. There are two or three synapses at most between the inner ear and the recorded cells. Looking back, it *would* have been very remarkable to find neurons working so close to the sensors with such delicate a nonlinearity as that of the matrix elements. If not for Yaskusheva et al's results, the most natural thing to imagine is that Purkinje cells would proceed to a linear combination of the sensory input, and this is the only thing I can positively report to have seen.

I shall conclude this vestibular-system-related part of my thesis by expressing the hope that the idea tested here might prove helpful in other regions further removed from the sensors (for instance, in the vestibular nuclei that receive projections from the cerebellum, or perhaps more likely in the thalamus). Given the numerical shortcomings of the approach described here, there would of course be much to be done on the computational side. And data would be needed: Matthieu Tihy's experiment was so nicely suited to our needs that I can only hope that it was not a unique opportunity to perform such tests as we did.

Bibliography

[1] M. Tihy, *Activité de la cellule de Purkinje au sein du système vestibulaire dans un contexte actif*, Ph. D. Thesis, Université Pierre et Marie Curie, 2016.

[2] D. Angelaki and K. Cullen, *Vestibular system: the many facets of a multimodal sense*Annu Rev Neurosci. **31** (2008), pp. 125-150.

[3] D. Bennequin and A. Berthoz, *Non-Linear Galilean vestibular receptive fields*, Conf Proc IEEE Eng Med Biol Soc (2011), pp.2273-2276.

[4] G. Marini, L. Provini and A. Rosina, *Gravity responses of Purkinje cells in the nodulus.* Exp. brain Res. **24** (1976), pp. 311-323.

[5] K. Cullen, *The neural encoding of self-motion*, Current Opinion in Neurobiology **21** (2011), pp. 1-9.

[6] K. Cullen, *The vestibular system: multimodal integration and encoding of self-motion for motor control*, Trends in Neurosciences **35(3)** (2012), pp. 185-196.

[7] H. Fushiki and N. Barmack, *Topography and reciprocal activity of cerebellar Purkinje cells in the uvula-nodulus modulated by vestibular stimulation*, Journal of Neurophysiology **78** (1997), 3083-3094.

[8] B. Sheliga, S. Yakushin, A. Silvers, T. Raphan and B. Cohen, *Control of spatial orientation of the angular vestibulo-ocular reflex by the nodulus and uvula of the vestibulocerebellum*, Ann. N. Y. Acad. Sci. **871** (1999), pp. 94-122.

[9] R. Fitzpatrick, J. Butler and B. Day, *Resolving head rotation for human bipedalism.* Current Biology **16** (2006), pp. 1509-14

[10] D. Angelaki, A. Shaikh, A. Green and J. Dickman, *Neurons compute internal models of the physical laws of motion*, Nature **430** (2004), pp. 560-564.

[11] T. Yakusheva, A. Shaikh, A. Green, P. Blazquez, J. Dickman and D. Angelaki, *Purkinje cells in posterior cerebellar vermis encode motion in an inertial reference frame*, Neuron **54(6)** (2007), pp. 973-985.

[12] T. Yakusheva, P. Blazquez and D. Angelaki, *Relationship between Complex and Simple Spike Activity in Macaque Caudal Vermis during Three-Dimensional Vestibular Stimulation*, The Journal of Neuroscience **30(24)** (2010), pp. 8111-8126.

Part III

On the analogy between the tempered dual of a reductive Lie group and that of its Cartan motion group

Chapter 7

How tempered representations of a semisimple Lie group contract to its Cartan motion group

Contents

Abstract

George W. Mackey suggested in 1975 that there should be analogies between the irreducible unitary representations of a noncompact semisimple Lie group G and those of its Cartan motion group − the semidirect product G_0 of a maximal compact subgroup of G and a vector space. In these notes, I focus on the carrier spaces for these representations and try to give a precise meaning to some of Mackey's remarks. I first describe a bijection, based on Mackey's suggestions, between the tempered dual of G − the set of equivalence classes of irreducible unitary representations which are weakly contained in $\mathbb{L}^2(G)$ − and the unitary dual of G_0. I then examine the relationship between the individual representations paired by this bijection: there is a natural continuous family of groups interpolating between G and G_0, and starting from the Hilbert space $\mathbf{H}$ for an irreducible representation of G, I prove that there is an essentially unique way of following a vector through the contraction from G to G_0 within a fixed Fréchet space that contains $\mathbf{H}$. It then turns out that there is a limit to this contraction process on vectors, and that the subspace of our Fréchet space thus obtained naturally carries an irreducible representation of G_0 whose equivalence class is that predicted by Mackey's analogy.

1 Introduction

When G is a Lie group and K is a closed subgroup of G, one can use the linear action of K on the vector space $V = Lie(G)/Lie(K)$ to define a new Lie group G_0 − the semidirect product $K \ltimes V$. This group is known as the contraction of G with respect to K, and it is famous in mathematical physics: the Poincaré group of special relativity admits as a contraction the Galilei group of classical inertial changes, and it is itself a contraction of the de Sitter group which appears in general relativity [1].

Since the unitary irreducible representations of the Poincaré group are well-known to yield particle state spaces for quantum physics, it is quite natural to wonder whether there is a relationship between the representation theories of G and G_0. For most Lie groups, including the Poincaré group, unitary representations do not behave well under the contraction: in general the parameters needed to identify representations of G and G_0 are rather different, and this is important for physics − a consequence of the bad behaviour in the case of the Poincaré group is that the notion of (rest) mass has different meanings in special and Galilean relativity.

In 1975 however, George Mackey − who had single-handedly developed the representation theory of semidirect products like G_0 in the 1950s − noticed that in the special case where G is a semisimple Lie group and K is a maximal compact subgroup of G, there is a coincidence between the parameters needed to describe rather large [2] subsets of the unitary duals of G and G_0, and some analogies in the way to build these subsets. In this case, G_0 is often called the Cartan motion group of G: it acts through rigid motions on the flat symmetric space G_0/K, while G is the isometry group of the negatively curved G/K.

It is rather surprising that there should be a deep analogy between the representation theories of these two groups, and not only because the algebraic structures of G and G_0 are very different. When Γ is a type I Lie group, let us write $\widehat{\Gamma}$ for its reduced (!) unitary dual − the set of equivalence classes of unitary irreducible representations which are weakly

1. See the lecture by Freeman Dyson [11].

2. In Mackey's suggestions, the meaning of "large" here refers to the Plancherel measure on the unitary dual of G_0.

contained in the regular representation; Mackey's theorems then make the description of $\widehat{G_0}$ quite accessible, while describing $\widehat{G}$ is a formidable task which took all of Harish-Chandra's talent and energy. Mackey nevertheless went on to conjecture that there should be a natural one-to-one correspondence between large enough subsets of $\widehat{G}$ and $\widehat{G_0}$:

> *In view of the facts outlined above for $SL_2(\mathbf{C})$ it is natural to wonder to what extent one can find a correspondence between "most" of the irreducible representations of G and those of the semidirect product G_0.*

The groups G and G_0 fit into a continuous one-parameter family (see Section 2), and in 1985 Dooley and Rice proved [8] that the operators for principal series representations of G do weakly converge, as the contraction is performed, to operators for a generic representation of G_0 (I discuss some of their results in section 4). Although the initial reactions to Mackey's ideas seem to have been rather skeptical, an interest in Mackey's suggestions later sprang from the deformation theory of $C^\star$-algebras: as Baum, Connes and Higson pointed out in [2, 6], the Baum-Connes conjecture for G (in its "smooth" version due to Connes and Kasparov, proved by Wasserman since then) is a precise counterpart to Mackey's analogy at the level of cohomology.

But the interest for Mackey's proposal seems to have waned since then, and it is scarcely – if at all – mentioned in the recent representation-theoretic literature (see however [4]). One of the reasons for the subject not having been pursued further, even after the mentioned developments in operator algebras, may be the fact that at the level of representation spaces, for the deeper strata of the unitary duals (as one moves away from the principal series of G or the generic representations of G_0), the analogy seems doomed to be rather poor: for instance, there are unitary irreducible representations of G_0 whose carrier spaces are finite-dimensional, while all unitary irreducible representations of G are infinite-dimensional (the trivial one excepted). Thus, as Mackey says,

> *Above all [the analogy] is a mere coincidence of parametrizations, with no evident relationship between the constructions of corresponding representations.*

However, Nigel Higson recently revived the subject [17, 18], starting back from the Connes-Kasparov conjecture and showing that for complex semisimple groups, a precise elaboration on Mackey's ideas leads to new proofs of the conjecture: the $C^\star$-algebra point of view involves shifting the attention from the representation spaces to their matrix coefficients, and Higson noticed that when G is complex semisimple, there is a deep, though not obvious, analogy between the structure of reduced $C^\star$-algebras of G and G_0: they turn out to be assembled from identical building blocks, and fit into a continous field [3] which turns out to be assembled from constant fields through Morita equivalences, extensions and direct limits. He also made the important side observation that while Mackey's suggestions treated the unitary dual as a Borel space, the K-theory in the Connes-Kasparov phenomenon treats it as a topological space, and there is no natural way to relate these two points of view. This suggests that Mackey's analogy should extend to the full tempered dual of G, yielding an interesting bijection between $\widehat{G_0}$ and all of $\widehat{G}$; this I will take up in Section 3 below.

3. The continuous field is defined using the deformation from G to G_0 which, from section 2 onwards, will play a key role in these notes.

In these notes, my aim is to define such a bijection when G is a real reductive Lie group, and to reconsider Mackey's proposal from a rather naive perspective: I will start with spaces realizing elements of $\widehat{G}$ and try to describe what happens to the (smooth, K-finite) vectors as one proceeds to the contraction. My hope is to give in this way a somewhat simpler picture than is usual for the relationship between $\widehat{G}$ and $\widehat{G}_0$, and to try to understand why the various strata in the unitary duals behave very differently.

As we shall see, the part of Mackey's analogy which relates spherical principal series representations of G to generic class-one representations of G_0 can be rephrased as transferring harmonic analysis on a symmetric space of the noncompact type to classical Fourier analysis on its (Euclidean) tangent space at a given point. This is a much-studied problem with beautiful ramifications [16, 35, 20, 10], and Higson's account of the Connes-Kasparov phenomenon shows that bringing Mackey's point of view into the picture is not at all devoid of interest in this case.

Now at the other end of the tempered spectrum, if G has discrete series representations, Mackey's proposal is that we should relate them to irreducible representations of K; what makes this reasonable is the fact that a discrete series representation has a unique minimal K-type in the sense of Vogan [41]. If G_0 is to be brought into the picture here, my task is then to understand how its minimal K-type can emerge from a discrete series representation as the contraction from G to G_0 is performed.

The methods I will use here are not original in any way: on the contrary, I will try to take full advantage of the geometric realizations of unitary representations of G which were set forth in the years immediately following Mackey's proposal. These realizations provide natural topologies, defined on (dense subsets of) the Hilbert spaces, that are different from the Hilbert space norm: at least for principal series and discrete series representations, the smooth vectors in the Hilbert spaces for tempered representations of G can be seen as functions on, or sections of homogeneous bundles on, the symmetric space G/K. In these notes I will trace Mackey's analogy to phenomena which are invisible to the Hilbert space topology, but become obvious when the topology of uniform convergence on compact subsets of G/K comes in.

Here is an outline of my notes. I will start with a description of the Cartan motion group, its unitary dual, and some aspects of the contraction from G to G_0 in section 2. After these preliminaries, I will write down a bijection between the tempered dual $\widehat{G}$ and $\widehat{G_0}$; for this extension of Higson's analysis in the complex case, the main tool is Vogan's minimal K-type parametrization of the tempered irreducible representations of G which have real infinitesimal character. From section 3.3 onwards I focus on individual representation spaces, pursuing evidence for a phenomenon described in section 3.3 and summarized as Theorem 3.2 there. Sections 4 and 5 examine what happens in the case of the spherical principal series representations and the discrete series representations, respectively: both start with a presentation of the geometric realizations I will use, then watch a vector evolve as one proceeds to the contraction from G to G_0. Since the full tempered dual of G can be in some sense assembled from the discrete series of reductive subgroups, I then use the results on the discrete series and the ideas (and lemmas) of the spherical principal case to work out the general case in sections 6 (real infinitesimal character) and 7 (general case).

Acknowledgments. I am very grateful to my doctoral advisor Daniel Bennequin, who encouraged my enthusiasm for this subject and offered constantly good advice. It is a pleasure to thank François Rouvière and David Vogan for very useful remarks and important corrections on an earlier version of this chapter, and Michel Duflo, Nigel Higson and Michèle Vergne for very helpful discussions.

2 The Cartan motion group and its unitary dual

2.1 The contraction from G to G_0

Throughout these notes, I shall consider a connected, linear, noncompact reductive Lie group G, assume (mainly for convenience) that it has compact center, and write $\mathfrak{g}$ for its Lie algebra. Let's start from a maximal compact subgroup K of G, its Lie algebra K, the Cartan involution θ of G with fixed-point-set K, and write $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{p}$ for the decomposition of $\mathfrak{g}$ into the $+1$ and -1 eigenspaces of the derivative of θ at 1_G. The adjoint action of K leaves $\mathfrak{p}$ invariant, so we can (re-)define G_0 as the set $K \times \mathfrak{p}$ with the group structure

$$(k_1, v_1) \cdot_0 (k_2, v_2) := (k_1 k_2, v_1 + \mathrm{Ad}(k_1) v_2).$$

Let me start this section by saying more precisely how G_0 is a deformation of G. For each $t > 0$, we can use the diffeomorphism

$$\begin{aligned} \varphi_t : K \times \mathfrak{p} &\to G \\ (k, v) &\mapsto \exp_G(tv)k \end{aligned}$$

to endow the set $K \times \mathfrak{p}$ with a group structure which turns φ_t into an isomorphism. So we will write G_t for be the set $K \times \mathfrak{p}$ with the composition

$$(k_1, v_1) \cdot_t (k_2, v_2) := \varphi_t^{-1}(\varphi_t(k_1, v_1) \cdot_G \varphi_t(k_2, v_2)).$$

Lemma 2.1. *For every k_1, k_2 in K and v_1, v_2 in $\mathfrak{p}$, the composition $(k_1, v_1) \cdot_t (k_2, v_2)$ goes to $(k_1, v_1) \cdot_0 (k_2, v_2)$ as t goes to zero.*

To prove it, let us write $(k_1, v_1) \cdot_t (k_2, v_2)$ as $(k(t), v(t))$ with $k(t)$ in K and $v(t)$ in $\mathfrak{p}$; we need to see that $k(t)$ goes to $k_1 k_2$ and $v(t)$ goes to $v_1 + \mathrm{Ad}(k_1)v_2$. The definition says

$$e^{tv(t)}k(t) = \left(e^{tv_1}k_1\right)\left(e^{tv_2}k_2\right) = e^{tv_1}e^{t\mathrm{Ad}(k_1)v_2}k_1k_2. \tag{2.1}$$

Now, recall that

$$u : \mathfrak{p} \xrightarrow{\exp_G} G \twoheadrightarrow G/K.$$

is a global diffeomorphism. Let me write $(g, x) \in G \times \mathfrak{p} \mapsto g \cdot x = ge^X K$ for the usual transitive action of G on G/K. From which we see that if $u : \mathfrak{p} \to G/K$ stands for the global diffeomorphism $X \mapsto e^X K$, then (2.1) says that $u[tv(t)] = e^{tv_1} \cdot u\,[t\mathrm{Ad}(k)v_2]$. As t goes to zero, we deduce that $u[tv(t)]$ goes to the origin $1K$ of G/K, so $tv(t)$ goes to zero in $\mathfrak{p}$, and then we obtain from (2.1) that $k(t)$ goes to k_1k_2.

As for the convergence of $v(t)$, applying the Cartan involutive automorphism of G whose fixed-point-set is K to both sides of (2.1) yields $e^{-tv(t)}k(t) = e^{-v_1}e^{-t\mathrm{Ad}(k_1 v_2}k_1k_2$; taking inverses and mutiplying with (2.1) in turn yields

$$e^{2tv(t)} = e^{tv_1}e^{2t\mathrm{Ad}(k_1)v_2}e^{tv_1}. \tag{2.2}$$

The Campbell-Hausdorff formula says that if t is small enough, then the right-hand-side can be written as $e^{2t(v_1 + \mathrm{Ad}(k_1)v_2 + r(t))}$, where $r(t)$ is an element of $\mathfrak{g}$ for each t (defined as the sum of a convergent Lie series in v_1 and $\mathrm{Ad}(k_1)v_2$) with $r(t) = O(t)$ (Landau notation). When t is close enough to zero, both $2tv(t)$ and $2t\,(v_1 + \mathrm{Ad}(k_1)v_2 + r(t))$ lie in a neighbourhood of zero in $\mathfrak{g}$ over which the restriction of $\exp_G$ is an injection, so (2.2) yields $v(t) = v_1 + \mathrm{Ad}(k_1)v_2 + r(t)$. We see that $r(t)$ in fact lies in $\mathfrak{p}$, and as t goes to zero we do obtain the convergence of $v(t)$ to $v_1 + \mathrm{Ad}(k_1)v_2$. $\square$

Before we delve into representation theory, I will record here the effect of this contraction on the Riemannian symmetric space G/K. Recall that we can start from a positive-definite quadratic form on the tangent space to G/K at the identity coset $\{K\}$ (for instance the one obtained from the restriction to $\mathfrak{p}$ of the Killing form), and then use the (left) action of G to build a G-invariant Riemannian metric on G/K; if the given scalar product form is appropriately normalized, this metric has constant scalar curvature -1.

2.2 The G_t-actions on $\mathfrak{p}$

2.2.1. Now, the Cartan decomposition provides an explicit diffeomorphism between G/K and $\mathfrak{p}$, so we can use it to make $\mathfrak{p}$ into a G-homogeneous space, and we can do this for each $t > 0$, using the natural maps

$$u_t : \mathfrak{p} \xrightarrow{\exp_{G_t}} G_t \twoheadrightarrow G_t/K.$$

The fact that these maps are diffeomorphisms provides us with a transitive action of G on $\mathfrak{p}$, for which I will write $(g, x) \in G \times \mathfrak{p} \mapsto g \cdot x$ in these notes, as well as a transitive action of each G_t, for which my notation will be $(\gamma, x) \in G_t \times \mathfrak{p} \mapsto \gamma \cdot_t x$. Of course the stabilizer of the point 0 in $\mathfrak{p}$ is K for each t. Let's also transfer the natural metric on G_t/K to $\mathfrak{p}$ through u_t, but take into account the fact that the Killing forms of G_t and G are not quite the same: so let us start from the scalar product on $\mathfrak{p}$ chosen at the end of the previous subsection, say B, and use it for each t to build a G_t-invariant metric η_t on $\mathfrak{p}$ which coincides with B at zero. Let me also write η for the G-invariant metric on $\mathfrak{p}$ built in this way and note that $\eta = \eta_1$.
If we do this, then the metric η_t has scalar curvature $-t^2$. On the other hand, we can build a G_0-invariant Euclidean metric η_0 on $\mathfrak{p}$ from B and the action of G_0 (which includes the translations of $\mathfrak{p}$, and for which my notation will be $(g_0, x) \in G_0 \times \mathfrak{p} \mapsto g_0 \cdot_0 x$), and the metrics η_t do tend to η_0 as t tends to zero (in the topology, say, of uniform convergence on compact sets for the metrics' coefficients in affine coordinates on $\mathfrak{p}$).

In the next three paragraphs, I will record very simple facts on this geometrical setting.

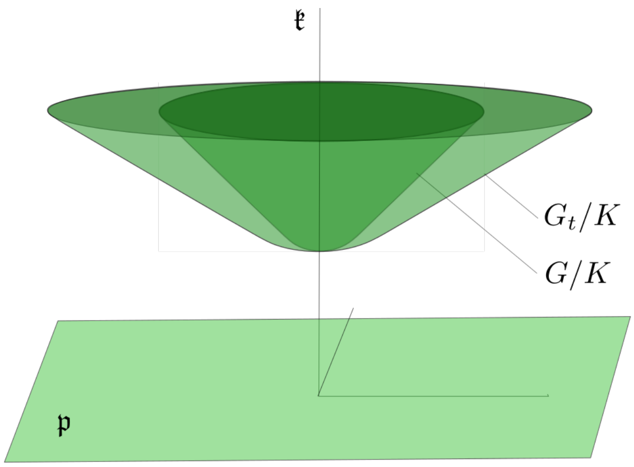


Figure 1: This is a picture of (co)adjoint orbits for G and G_t when G is $SL_2(\mathbb{R})$. The (green) horizontal plane is the space $\mathfrak{p}$ of symmetric matrices with zero trace, the vertical axis is the line $\mathfrak{k}$ of antisymmetric matrices; the drawn orbits are the G-and G_t-coadjoint orbits of their common intersection with the vertical axis. Section 2.2.4 is a comment on this figure (see especially lemma 2.4).

2.2.2. Let me consider the dilation

$$\begin{aligned} z_t \ &: \mathfrak{p} \to \mathfrak{p} \\ x &\mapsto \frac{x}{t}. \end{aligned}$$

An important ingredient in these notes will be the fact that the relationship between G and G_t is simple enough that their actions on $\mathfrak{p}$ are related through z_t:

Lemma 2.2. *For every x in $\mathfrak{p}$ and every g in G, $\varphi_t^{-1}(g) \cdot_t z_t(x)$ is equal to $z_t\,(g \cdot x)$.*

Proof. Let us see where the diffeomorphism u_t sends the both of them. On the one hand,

$$u_t\left(\varphi_t^{-1}(g) \cdot_t z_t(x)\right) = \varphi_t^{-1}(g)\exp_{G_t}(z_t x)K = \varphi_t^{-1}(g\exp_G(x)K),$$

and on the other hand, using the definition of the G_t and G-actions,

$$u_t\,(z_t\,(g \cdot x)) = \exp_{G_t}[z_t(g \cdot x)]\,K = \varphi_t^{-1}(\exp_G(g \cdot x)K) = \varphi_t^{-1}\,(g\exp_G(x)K)\,.$$

$\square$

2.2.3. Here is a remark which says how the action of G_t on $\mathfrak{p}$ admits the natural action of G_0 as a limit. With a mind to make the G_t-actions more easily legible, let me introduce the diffeomorphisms corresponding to the Cartan decomposition of each G_t, writing

$$\begin{aligned} \alpha_t : G_0 &\to G_t \\ (k, v) &\mapsto \exp_{G_t}(v)k. \end{aligned}$$

and $\alpha : G_0 \to G, (k, v) \mapsto \exp_G(v)k$.

When we remember that G_0 and G_t coincide as sets, α_t simply is the identity, so $\alpha_t(k, v)$ should be thought of as shorthand for "(k, v) viewed as an element of G_t". Note that none of the α_t is a group morphism, and that φ_t sends $\alpha_t(k, v)$ to $\alpha(k, tv)$.

Lemma 2.3. *For each g_0 in G_0, $\alpha_t(g_0) \cdot_t x$ tends to $g_0 \cdot_0 x$ uniformly on compact subsets of $\mathfrak{p}$ as t goes to zero.*

Proof. Recall that the G_t-action is defined by using

$$u_t : \mathfrak{p} \xrightarrow{\exp_{G_t}} G_t \twoheadrightarrow G_t/K.$$

to transfer the action (by left multiplication) of G_t on G_t/K.

Now if $g_0 = (k, v)$,

$$\begin{aligned} \alpha_t(g_0) \cdot_t x &= \varphi_t^{-1}\,[\alpha(k, tv)] \cdot_t x \\ &= \frac{1}{t}\,[\alpha(k, tv)] \cdot (tx) \quad \text{by Lemma 2.1.} \end{aligned}$$

What we need to show is that $\frac{1}{t}\alpha(k, tv) \cdot (tx)$ goes to $v + \mathrm{Ad}(k)x$ as t goes to zero. But using the diffeomorphism u between G and G/K, we have

$$
\begin{aligned}
u\left(\frac{1}{t}\alpha(k,tv)\cdot(tx)\right) &= \exp_G\left(\frac{1}{t}u^{-1}\left[\alpha(k,tv)\exp_G(tx)K\right]\right)K \\
&= \exp_G\left(\frac{1}{t}u^{-1}\left[\exp_G(tv)k\exp_G(tx)K\right]\right)K \\
&= \exp_G\left(\frac{1}{t}u^{-1}\left[\exp_G(tv)\exp_G(t\mathrm{Ad}(k)[x])K\right]\right)K.
\end{aligned}
$$

Now, the discussion in section 2.1 shows that $e^{tv}e^{t\mathrm{Ad}(k)[x])}$ can be written as $e^{t\beta(t)}\kappa(t)$, with $\kappa(t)$ in K and $\beta(t)$ in $\mathfrak{p}$, with $\beta(t) = v + \mathrm{Ad}(k)[x] + r(t)$, $r(t)$ in $\mathfrak{p}$, $r(t) = O(t)$. Then $\exp_G(tv)\exp_G(t\mathrm{Ad}(k)[x])K$ is equal to $\exp_G(t\beta(t))K = u\,[t\beta(t)]$, and

$$
u\left(\frac{1}{t}\alpha(k,tv)\cdot(tx)\right) = \exp_G\left(\frac{1}{t}\left[t\beta(t)\right]\right)K = u\left(v + \mathrm{Ad}(k)[x] + r(t)\right),
$$

and this does go to $v + \mathrm{Ad}(k)x$ as t goes to zero. □

2.2.4. The simple facts I just described are well-displayed by Figure 1. Whenever G/K is a hermitian symmetric space, which is the case when G is $SL_2(\mathbb{R})$, there is an element in $\mathfrak{k} \subset \mathfrak{g}$ whose stabilizer under the adjoint action of G is K. Because $\mathfrak{g}$ is the underlying vector space for the Lie algebra of G_t, too, we can look at the relationship between the orbits (although the adjoint and coadjoint actions are naturally equivalent for semisimple Lie groups, the usual vector space identification between $\mathfrak{g}_t$ and $\mathfrak{g}_t^\star$ depends on the Killing form, which varies with t: while for some questions − like studying the distribution characters − it would be desirable to look at coadjoint orbits of course, I will stick with adjoint orbits here).

Let me write $\phi_t : \mathfrak{k}\oplus\mathfrak{p} \to \mathfrak{k}\oplus\mathfrak{p}$ for the derivative $X_\mathfrak{k} + X_\mathfrak{p} \mapsto X_\mathfrak{k} + tX_\mathfrak{p}$ of φ_t at the identity − ϕ_t is a map from $\mathfrak{g}$ to itself.

Lemma 2.4. *The image under ϕ_t of a G_t-adjoint orbit, say Ω, is a G-adjoint orbit which has the same intersection with $\mathfrak{k}$ as Ω.*

Proof. Start from the fact that φ_t is a group morphism, so $\varphi_t^{-1}(g)\varphi_t^{-1}(h)\varphi_t^{-1}(g) = \varphi_t^{-1}(ghg^{-1})$ for $g, h \in G$, and just take the derivative at at $h = \mathbf{1}_G$; this yields

$$
Ad_{G_t}(\varphi_t^{-1}g)\left[\phi_t^{-1}\lambda\right] = \phi_t^{-1}\mathrm{Ad}_G(g)[\lambda]
$$

so the G-adjoint orbit for λ is the image under ϕ_t of the G_t-adjoint orbit for $\phi_t^{-1}\lambda$, as announced. □

On the face of it Figure 1 seems to be a complete picture of the contents of this subsection when G is $SL_2(\mathbb{R})$, with the metrics η_t transferred from those of the upper sheets of various hyperboloids to $\mathfrak{p}$ through the vertical projection. It would be a really complete picture if $u_t : \mathfrak{p} \to G_t/K$ were the vertical projection. This is not the case, but it is a near miss: an explicit calculation shows that u_t^{-1} is not quite the vertical projection on $\mathfrak{p}$, but that there is a very simple diffeomorphism of $\mathfrak{p}$, namely

$$
\tau : x \mapsto \frac{\sinh(\|x\|_B)}{\|x\|_B}(\mathcal{I}\ x)
$$

(where $\mathcal{I}$ is a rotation of angle $\pi/2$ and $\|\cdot\|_B$ is the norm induced by B), such that $u \circ \tau : \mathfrak{p} \to G/K$ coincides with the vertical projection.

The appearance of τ is not very surprising here [4]: the geodesics of G/K which go through the identity coset are sent by both u_t and the vertical projection to straight lines through the origin in $\mathfrak{p}$, and the (nonconstant) dilation factor in τ compensates for the difference between the speeds at which geodesics spread in hyperbolic space and Euclidean space (see [3], Chap. 6).

2.2.5. In these notes, I shall try to trace the relationship between the representations of G and G_0 to the fact that the building blocks of the tempered representation theory of G amount to studying spaces of functions on G/K, or sections of homogeneous bundles on it, which satisfy some invariant partial differential equation. The diffeomorphisms u_t, the relationship between the above actions of G and G_t on $\mathfrak{p}$ and between the metrics η_t, will be simple enough to allow me to follow vectors through the contraction. But before I focus on individual representations from section 3.3 onwards, I will set to describe a common parametrization for $\widehat{G}$ and $\widehat{G_0}$.

2.3 The unitary dual of G_0

Let me start with some standard notations: suppose $\mathfrak{a}$ is a maximal abelian subalgebra of $\mathfrak{p}$, and W is the Weyl group of the pair $(\mathfrak{g}_{\mathbf{C}}, \mathfrak{a}_{\mathbf{C}})$ (here and throughout these notes, a subscript $\mathbf{C}$ codes for the complexification of a Lie algebra). When χ is an element of $\mathfrak{p}^\star$ (the dual vector space), I write K_χ for its stabilizer (or "isotropy group") under the coadjoint action of K on $\mathfrak{p}$; note that K_χ is usually not connected, though as we shall see it has finitely many components. One can use the Killing form of $\mathfrak{g}$ to embed $\mathfrak{a}^\star$ in $\mathfrak{p}^\star$ as those linear forms on $\mathfrak{p}$ which vanish on the orthocomplement of $\mathfrak{a}$, and then note that all χs in $\mathfrak{a}^\star$ which are regular (that is, whose K-orbit in $\mathfrak{p}^\star$ has the largest possible dimension for a K-orbit in $\mathfrak{p}^\star$) have the same stabilizer; I will write M for it. I also write A for the abelian subgroup $\exp_G(\mathfrak{a})$.

I will now describe Mackey's results on the unitary dual of G_0 [28, 29].

Definition. *A* Mackey datum *is a couple* (χ, μ) *in which* χ *is an element of* $\mathfrak{a}^\star$*, and* μ *is an equivalence class of irreducible* K_χ*-modules.*

This non-standard vocabulary will be useful for us later; but each K-orbit in $\mathfrak{p}^\star$ intersects $\mathfrak{a}^\star$, so choosing a Mackey datum is the same as choosing first a K-orbit in $\mathfrak{p}^\star$, then an irreducible representation of the isotropy group of one of its elements − these are the more usual parameters for $\widehat{G_0}$.

From a Mackey datum $\delta = (\chi, \mu)$, one can produce a unitary representation of G_0 by unitary induction: set

$$\mathbf{M}_0(\delta) := \operatorname{Ind}_{K_\chi \ltimes \mathfrak{p}}^{G_0} \left[\mu \otimes e^{i\chi}\right].$$

If we write W for an irreducible K_χ-module of class μ, a Hilbert space for $\mathbf{M}_0(\delta)$ is obtained by considering

$$\left\{f : K \to W \mid f(km) = \mu(m)^{-1} f(k) \text{ for } k \in K, m \in K_\chi\right\}, \tag{2.3}$$

4. Thank you to Martin Puchol for pointing this out.

declaring that

$$g = (k, v) \in G_0 \text{ acts through } f \mapsto \left[u \mapsto e^{i\langle \chi, Ad(u^{-1})v\rangle} f(k^{-1}u)\right], \tag{2.4}$$

and for the Hilbert space structure taking the $\mathbf{L}^2$ space associated to the Haar measure of K and a K_χ-invariant inner product on W.

Mackey proved ([28, 29]) that each of the $\mathbf{M}_0(\delta)$ is irreducible. Moreover, if we start from two Mackey data $\delta_1 = \mathbf{M}(\chi_1, \mu_1)$ and $\delta_2 = \mathbf{M}(\chi_2, \mu_2)$, the condition for $\mathbf{M}_0(\delta_1)$ and $\mathbf{M}_0(\delta_2)$ to be unitarily equivalent is that there be an element of the Weyl group of $(\mathfrak{g}_\mathbb{C}, \mathfrak{a}_\mathbb{C})$ which sends χ_1 to χ_2 and μ_1 to an irreducible K_{χ_2}-module which is unitarily equivalent with μ_2. So we have an equivalence relation between Mackey data, and an injective map from the set of equivalence classes of Mackey data into $\widehat{G_0}$. Mackey also proved that this map is surjective: the assignment

$$\delta \mapsto \mathbf{M}_0(\delta)$$

gives a bijection between the Mackey data, up to equivalence, and the unitary irreducible representations of G_0, up to unitary equivalence (see the proofs I gave in Chapter 5).

Let me insist that this parametrization gives a stratification of $\widehat{G_0}$ into subsets recording the dimension of the orbit of the parameter χ in $\mathfrak{p}^\star$; parameters for the extreme strata are :

• Mackey data with $\chi = 0$, which correspond to the irreducible representations of K, with finite-dimensional carrier spaces;

• Mackey data with regular χ; the corresponding representations of G_0 are unitarily induced from $M \ltimes \mathfrak{p}$, so they have a realization as spaces of square-integrable vector-valued functions on K/M which transform according to (2.4) under the action of G_0. There is a simple geometric picture for [5] $\mathbf{M}_0(\chi, 1)$ which will be useful for us in section 4: consider the tempered distributions on $\mathfrak{p}$ whose Euclidean Fourier transform − a tempered distribution on $\mathfrak{p}^\star$ − is supported on $\mathrm{Ad}^\star(K) \cdot \chi$. These are automatically smooth functions on $\mathfrak{p}$, and when realized as functions on K/M through the Fourier transform, they do transform in the right way: (2.4) can be easily understood from the usual formula for the Fourier transform of $x \mapsto f(g_0 \cdot_0 x)$. By considering the smooth and square-integrable functions whose Fourier transform has the mentioned property, we get a realization of $\mathbf{M}_0(\chi, 1)$ for which members of the carrier space appear as functions on $\mathfrak{p}$ which are combination of those plane waves with frequency vectors on $\mathrm{Ad}^\star(K) \cdot \chi$.

To close this subsection, let me note that every unitary irreducible representation of G_0 is weakly contained in the regular representation, so the reduced dual and the unitary dual of G_0 coincide. This is in sharp contrast with the situation for our reductive group G, for which the unitary dual is quite larger than the reduced dual; to give the simplest but significant example, the trivial representation of G is not in the reduced dual of G. In fact, the unitary irreducible representations which appear in the reduced dual of G are all *tempered* − this means that their matrix coefficients lie in $\mathbf{L}^{2+\epsilon}(G)$ for each positive ϵ, and although this definition leaves the terminology rather mysterious, it makes it quite clear that the trivial representation is not tempered.

When we write out a correspondence between $\widehat{G_0}$ and $\widehat{G}$ in the next section, the trivial representation of G_0 will thus be associated with a quite non-trivial (and infinite-dimensional) representation of G.

5. The 1 here means that I use the trivial representation of K_χ.

3 Mackey's correspondence

3.1 Minimal K-type for discrete series, and a theorem of Vogan

In this subsection I will assume that G has a nonempty discrete series and write T for a Cartan subgroup of K, so that T is also a (compact) Cartan subgroup of G.

Let us start with a unitary irreducible representation π of G in a Hilbert space $\mathbf{H}$; the restriction $\pi\big|_K$ is a direct sum of irreducibles.

Given the choice of a system Δ_c^+ of positive roots for the pair $(\mathfrak{k}_\mathbf{C}, \mathfrak{t}_\mathbf{C})$, let's write ρ_c for the half-sum of the elements of Δ_c^+; it is an element of $i\mathfrak{t}^\star$. Recall that an element λ of $\widehat{K}$ then has a *highest weight*, which is an element of $i\mathfrak{t}^\star$; I shall also write λ for it. An element of $\widehat{K}$ is a *minimal K-type* of π when, among the positive numbers $\|\lambda' + 2\rho_c\|$ in which λ' is the highest weight of a class occuring in $\pi|_K$, $\|\lambda + 2\rho_c\|$ is minimal (here $\|\cdot\|$ means the norm induced by the Killing form). We shall need only very simple instances of the deep problem of studying the minimal K-types in a unitary irreducible representation [41, 42].

The starting point for our common parametrization of $\widehat{G}$ and $\widehat{G_0}$ is the fact that a *discrete series* representations π of G has a *unique* minimal K-type, and that non-equivalent discrete series representations have non-equivalent minimal K-types. This was conjectured by Blattner in the wake of Harish-Chandra's formidable work on the discrete series, and proved by Hecht and Schmid; see [9, 13] and the historical remarks in [22].

Let's record here that later in these notes, it will also be an important fact that its minimal K-type in fact occurs with multiplicity one in $\pi|_K$.

This theorem gives a very precise indication of which representations of G we should attach to the subset of $\widehat{G}_0$ gathering the representations of K. But not all K-types are to be obtained as minimal K-types of discrete series representations of G, and in the unitary dual of G_0, there is no difference to be made between the various classes in $\widehat{K}$; to cover the remaining cases we will need a theorem of David Vogan which identifies tempered irreducible representations of G with a unique minimal K-type, in a way which treats all the elements in $\widehat{K}$ on the same footing.

For the rest this subsection, I no longer assume G to be connected or to have compact center, but only that it is a linear reductive Lie group in Harish-Chandra's class with abelian compact subgroups (see Chapter 0, §1 in [42]) − the induction steps in the next subsection will make this technical detail necessary.

To state Vogan's theorem, let's first recall the notion of infinitesimal character [6] (this presentation is taken from [39]).

When we consider the infinitesimal counterpart to an irreducible representation of G − a representation of the Lie algebra $\mathfrak{g}$ on a space V, and its complexification, the elements in the center of the universal enveloping algebra $\mathfrak{Z}(\mathfrak{g}_\mathbf{C})$ act as scalar multiples of the identity on V; we obtain an abelian character of the algebra $\mathfrak{Z}(\mathfrak{g}_\mathbf{C})$.

Now whenever $\mathfrak{h}_\mathbf{C}$ is a Cartan subalgebra of $\mathfrak{g}_\mathbf{C}$ and $W_\mathfrak{h}$ is the corresponding Weyl group, there is a simple correspondence between characters of the commutative algebra $\mathfrak{Z}(\mathfrak{g}_\mathbf{C})$ on the one hand, and $\mathfrak{h}_\mathbf{C}^\star/W_\mathfrak{h}$ on the other hand.

To define it, recall that Harish-Chandra defined an isomorphism $\xi_\mathfrak{h}$ from $\mathfrak{Z}(\mathfrak{g}_\mathbf{C})$ to the set $\mathbf{S}(\mathfrak{h}_\mathbf{C})^{W_\mathfrak{h}}$ of $W_\mathfrak{h}$-invariant symmetric polynomials on $\mathfrak{h}_\mathbf{C}$ (for the definition of $\xi_\mathfrak{h}$, see

6. It is not only for the reader's convenience that I recall the definition here: I will need it in Section 6.1 below.

the proof of lemma 6.5 below). Since evaluation at an element $\lambda \in \mathfrak{h}^\star_{\mathbf{C}}$ yields a map from $\mathbf{S}(\mathfrak{h}_{\mathbf{C}})$ to $\mathbf{C}$, we can compose with the Harish-Chandra isomorphism to obtain an abelian character, say $\xi_{\mathfrak{h}}(\lambda)$, of $\mathfrak{Z}(\mathfrak{g}_{\mathbf{C}})$. Once the obvious equivalences are quotiented out, the map $\lambda \mapsto \xi_{\mathfrak{h}}(\lambda)$ provides us with the promised bijection between $\mathfrak{h}^\star_{\mathbf{C}}/W_{\mathfrak{h}}$ and the set of abelian characters of $\mathfrak{Z}(\mathfrak{g}_{\mathbf{C}})$.

So if we start with an element λ in $\mathfrak{h}^\star_{\mathbf{C}}$, we see what it means for an irreducible representation π of G to have infinitesimal character λ. To state Vogan's theorem we need to see what it means for π to have real infinitesimal character. Assume $\mathfrak{h}_{\mathbf{C}}$ is the complexification of the Lie algebra $\mathfrak{h}$ of a Cartan subgroup of G which is stable under the Cartan involution associated to K. Splitting the subgroup into compact and vector subgroups, we get a decomposition $\mathfrak{h} = \mathfrak{t} \oplus \mathfrak{a}$, and we can set $\mathfrak{R}(\mathfrak{h}) = i\mathfrak{t} \oplus \mathfrak{a}$. This is a subspace of $\mathfrak{h}_{\mathbf{C}}$, and of course it is a real form of $\mathfrak{h}_C$; so our linear functional λ reads $\mathfrak{R}(\lambda) + i\mathfrak{I}(\lambda)$, where $\mathfrak{R}(\lambda)$ and $\mathfrak{I}(\lambda)$ are elements of $\mathfrak{h}^\star_{\mathbf{C}}$ whose restriction to $\mathfrak{R}(\mathfrak{h})$ is real-valued.

The phrase "π has real infinitesimal character", i.e. $\mathfrak{I}(\lambda) = 0$, then turns out to be independent of the choice of H and of the chosen representative λ of the infinitesimal character.

For $SL_2(R)$, there are but three irreducible tempered representations with real infinitesimal character which are not in the discrete series: the two "limits of discrete series" (which can be realized, like the discrete series representations, as spaces of holomorphic or anti-holomorphic functions on the hyperbolic plane, but have a rather different Hilbert space structure) and the "principal series representations with spectral parameter zero" (the name and properties of this one are different, see section 6.3). Each of these three turns out to have a unique minimal $SO(2)$-type; the corresponding characters are the trivial one, the "identity" character corresponding to the usual embedding of $SO(2)$ in $\mathbf{C}$, and the "conjugation" character (the complex-conjugate of the former); they are precisely the characters that do not appear as minimal $SO(2)$-type in any discrete series representation.

Returning to a linear reductive Lie group G in Harish-Chandra's class with abelian Cartan subgroups, Vogan proved that every irreducible tempered representation of G which has real infinitesimal character has a unique minimal K-type, that nonequivalent such representations have different minimal K-types, and that all K-types can be obtained in this way. This can be rephrased as follows (see for instance Theorem 1.2 in [40]) .

Theorem (Vogan). *The minimal K-type map defines a bijection between the equivalence classes of irreducible tempered representations of G which have real infinitesimal character, and the equivalence classes of irreducible representations of K.*

Our results in sections 5 and 6 can be viewed as a way to use the contraction from G to G_0 to exhibit its minimal K-type from the carrier space of a tempered irreducible representation with real infinitesimal character. But now let us linger at the level of parameters; it is time to give a precise meaning to Mackey's analogy.

3.2 A bijection between the tempered duals

Let us come back to a linear connected reductive group G whose center is compact. I am now going to define a map from $\widehat{G_0}$ to $\widehat{G}$. Let us start with a Mackey datum $\delta = (\chi, \mu)$.

Out of $\chi \in \mathfrak{a}^\star$, we first build a parabolic subgroup P_χ of G. Consider the centralizer L_χ of χ in G (for the coadjoint action of G on $\mathfrak{g}^\star$). Then L_χ is the Cartan-involution-stable Levi factor of a parabolic subgroup of G (see [39], Lemma 3.4(4)). Writing $L_\chi = M_\chi A_\chi$

for the Langlands decomposition of L_χ, there is a connected nilpotent subgroup N_χ of G such that

$$P_\chi = L_\chi N_\chi = M_\chi A_\chi N_\chi$$

is a parabolic subgroup of G.

Here M_χ is a reductive subgroup of G that contains M, A_χ is a subgroup of A (so the Lie algebra $\mathfrak{a}_\chi$ is a subspace of $\mathfrak{a}$, and A_χ is $\exp_G(\mathfrak{a}_\chi)$), and M_χ centralizes A_χ. The subgroup M_χ contains K_χ as a maximal compact subgroup by definition. It is no longer semisimple nor connected in general; however, it is a reductive linear group in Harish-Chandra's class and it does have abelian Cartan subgroups.

Note that L_0 is all of G, while $L_\chi = MA$ whenever χ is regular.

For arbitrary χ, it is worth recalling that L_χ is generated by M and by the root subgroups for the roots of $(\mathfrak{g}, \mathfrak{a})$ whose scalar product with χ is zero; in addition, A_χ is then by definition the intersection of the kernels of these same roots. As for the definition of N_χ, recall that an ordering of $\mathfrak{a}^\star$ defines a set of positive weights for $\mathfrak{a}_\chi$ in $\mathfrak{g}$; the sum of positive weight spaces then yields a subalgebra $\mathfrak{n}_\chi$ of $\mathfrak{g}$, and setting $N_\chi = \exp_G(\mathfrak{n}_\chi)$ makes $L_\chi N_\chi$ a parabolic subgroup.

⋆

We now perform parabolic induction from P_χ. Our Mackey datum δ came with $\mu \in \widehat{K_\chi}$; since M_χ admits K_χ as a maximal compact subgroup and is a linear reductive group in Harish-Chandra's class with abelian Cartan subgroups, we can use Vogan's theorem to attach to μ the irreducible tempered representation $\mathbf{V}_{M_\chi}(\mu)$ of M_χ. We extend χ to define a one-dimensional representation of $A_\chi N_\chi$, and then consider the unitarily induced representation

$$\mathbf{M}(\delta) := \mathrm{Ind}_{P_\chi}^{G}\left[\mathbf{V}_{M_\chi}(\mu) \otimes e^{i\chi}\right].$$

There are important results of representation theory to be called upon here. They are simple consequences of deep work on irreducible tempered representations by Harish-Chandra on the one hand, Knapp and Zuckerman on the other, but since I will need to check a few details let me state them as three lemmas :

Lemma 3.1. *For each Mackey datum δ, this $\mathbf{M}(\delta)$ is irreducible and tempered.*

Lemma 3.2. *Suppose δ_1, δ_2 are Mackey data. Then the representations $\mathbf{M}(\delta_1)$ and $\mathbf{M}(\delta_2)$ are unitarily equivalent if and only if δ_1 and δ_2 are equivalent as Mackey data.*

The results on $\widehat{G_0}$ recalled above mean that we get an injection from $\widehat{G_0}$ into $\widehat{G}$. Now, here is a consequence of Knapp and Zuckerman's work.

Lemma 3.3. *Every irreducible tempered unitary representation is equivalent with one of the representations $\mathbf{M}(\delta)$.*

These three lemmas can be summarized by the following result.

Theorem 3.1. *The correspondence $\mathbf{M} \circ \mathbf{M}_0^{-1}$ induces a bijection between the tempered dual $\widehat{G}$ and the unitary dual $\widehat{G_0}$.*

To prove this theorem, we need only relate our three lemmas to an irreducibility theorem by Harish-Chandra on the one hand, and to the Knapp-Zuckerman classification of tempered irreducible representations on the other hand.

Proof of Lemma 3.1 and Lemma 3.2. I will use a result of Harish-Chandra, cited as theorem 14.93 in [22] (for a full proof and discussion see [26], theorem 4.11) − this also appears in [39], Lemma 3.2(5). Say that an element of $i\mathfrak{a}_\chi^\star$ is $\mathfrak{a}_\chi$*-regular* when its scalar product with each root of $(\mathfrak{g}_{\mathbf{C}}, \mathfrak{a}_{\chi,\mathbf{C}})$ (every nonzero weight of the adjoint representation of $\mathfrak{a}_{\chi,\mathbf{C}}$ on $\mathfrak{g}_{\mathbf{C}}$) is nonzero. The result by Harish-Chandra implies Lemma 3.1 if we can ensure that χ is $\mathfrak{a}_\chi$-regular.

Recall that a root of $(\mathfrak{g}_{\mathbf{C}}, \mathfrak{a}_{\chi,\mathbf{C}})$ is a root of $(\mathfrak{g}_{\mathbf{C}}, \mathfrak{a}_{\mathbf{C}})$ that does not vanish on $\mathfrak{a}_\chi$. Now, if β is a root of $(\mathfrak{g}_{\mathbf{C}}, \mathfrak{a}_{\mathbf{C}})$ whose scalar product with χ is zero, then it is a root of $(\mathfrak{l}_{\chi,\mathbf{C}}, \mathfrak{a}_{\mathbf{C}})$, and as recalled above, $\mathfrak{a}_\chi$ is then contained in the kernel of β; as a consequence, β must vanish on $\mathfrak{a}_\chi$, and β cannot be a root of $(\mathfrak{g}_{\mathbf{C}}, \mathfrak{a}_{\chi,\mathbf{C}})$.

So it is indeed true that χ is $\mathfrak{a}_\chi$-regular, and can conclude that for each Mackey datum δ, $\mathbf{M}(\delta)$ satisfies the hypotheses of [22], theorem 4.93, from which Lemma 3.1 follows (the "temperedness" part is standard; see e.g. [22], Chapter VII, § 10-11). Now that we know that the hypothesis in Harish-Chandra's criterion as stated in [26], theorem 4.11, is satisfied, *Lemma 3.2* follows from it as well. □

Proof of Lemma 3.3. This is a straightforward consequence of Knapp and Zuckerman's classification of irreducible tempered representations. Suppose π is an irreducible tempered representation of G. Then there is a parabolic subgroup $P = M_P A_P N_P$ of G (usually not minimal), a tempered representation δ of M_P with real infinitesimal character and an element ν of $(\mathfrak{a}_P)^\star$, such that π is equivalent with

$$\mathrm{Ind}_{M_P A_P N_P}^{G}\left(\delta \otimes e^{i\nu}\right)$$

(see for instance [39], Theorem 3.3, and of course Theorem 16.10 in [22], along with [24]). What we need to prove is that this is in the image of our Mackey map $\mathbf{M}$. A first remark is that P is contained in P_ν: because ν is an element of $\mathfrak{a}_P$, the centralizer $L_P = M_P A_P$ of $\mathfrak{a}_P^\star$ (for the coadjoint action) is contained in that of ν, so L_P is contained in L_ν. This implies that P is contained in P_ν, that M_ν contains M_P, A_ν is contained in A_P and N_ν is contained in N_P.

Let's introduce the subgroups $\tilde{A}, \tilde{N}$ whose Lie algebras are the orthocomplements of $\mathfrak{a}_\nu$ and $\mathfrak{n}_\nu$ in $\mathfrak{a}_P$ and $\mathfrak{n}_P$, respectively; then $A_P = A_\nu \tilde{A}$, $N_P = N_\nu \tilde{N}$ and because $\mathfrak{m}_\nu$ is orthogonal to $\mathfrak{a}_\nu \oplus \mathfrak{n}_\nu$, both $\tilde{A}$ and $\tilde{N}$ are contained in M_ν. We can then use the fact that A_P is abelian and N_P normalizes A to write

$$\mathrm{Ind}_{M_P A_P N_P}^{G}\left(\tau \otimes e^{i\nu} \otimes 1\right) = \mathrm{Ind}_{(M_P \tilde{A}\tilde{N})A_\nu N_\nu}\left((\tau \otimes e^0) \otimes e^{i\nu}\right).$$

Then $\tilde{P} = M_P \tilde{A}\tilde{N}$ is a subgroup of M_ν, $M\tilde{A}$ is the centralizer of $\tilde{A}$ in M_ν: so $\tilde{P}$ is in fact a parabolic subgroup of M_ν. Now, $\sigma = \mathrm{Ind}_{M\tilde{A}\tilde{N}}^{M_\nu}(\delta \otimes e^0)$ is a tempered representation of M_ν, it has real infinitesimal character, and it is irreducible (otherwise π would not be !). The double induction formula yields

$$\mathrm{Ind}_{P_\nu}^{G}\left(\sigma \otimes e^{i\nu}\right) = Ind_{M_\nu A_\nu N_\nu}^{G}\left(\mathrm{Ind}_{\tilde{P}}^{M_\nu}\left(\tau \otimes e^0\right) \otimes e^{i\nu}\right)$$

which proves that π is in the image of $\mathbf{M}$; this is lemma 3.3. □

Remark. The proof of Lemma 3.3 uses a "weak" version of Knapp and Zuckerman's classification, but I will later need to refer to a more precise form of the result; see section 6. To prepare the way for sections 3.2 and 6, let me say here that Knapp and Zuckerman actually proved that if π is an irreducible tempered representation of G, then there is a *cuspidal* parabolic subgroup $P = M_P A_P N_P$ of G (one in which M_P has a nonempty discrete series), a *discrete series or nondegenerate limit of discrete series* representation δ of M_P, and an element ν of $(\mathfrak{a}_P)^\star$, such that π is equivalent with $\mathrm{Ind}_{M_P A_P N_P}^{G}\left(\delta \otimes e^{i\nu}\right)$. For the definition of limits of discrete series and other consequences of this, see section 6 below.

Remark. It can be proven that the bijection in Theorem 3.1 is compatible with Vogan's classification of tempered representations by minimal K-types: the set of minimal K-types in an irreducible tempered representation of G, even if there is more than one element in it, is the same as that of the corresponding representation of G_0. The result would be rather out of place in this chapter, so it will appear in the next chapter (chapter 8, Proposition 2.1).

3.3 Program for sections 4 to 7

If we start with a Mackey datum δ, the constructions above provide a Hilbert space $\mathbf{H}$ and a morphism π from G to the unitary group of $\mathbf{H}$; but we can also view δ as a Mackey datum for each of the G_t, getting a Hilbert space $\mathbf{H}_t$ and a morphism $\pi_t : G_t \to \mathbf{U}(\mathbf{H}_t)$ for each $t > 0$. Now we have an explicit isomorphism φ_t from G_t to G, and the morphisms $\pi_t \circ \varphi_t^{-1}$ and π define irreducible representations of G. If we are careful about the interpretation of δ as a Mackey datum for G_t (see sections 4.2 and 7.2), they will be unitarily equivalent. In this case I will say that the equivalence class of π_t is $\mathbf{M}_t(\delta)$.

Definition. *Suppose π is a unitary representation of G with Hilbert space $\mathbf{H}$ and π_t is a unitary representation of G_t with Hilbert space $\mathbf{H}_t$. A linear map*

$$\mathbf{C}_t : \mathbf{H} \to \mathbf{H}_t$$

will be called a contraction map *when it intertwines π and $\pi_t \circ \varphi_t^{-1}$.*

Notice that Schur's lemma says there cannot be many contraction maps; when there is one it is unique up to to a scalar of modulus one, and as we shall see, upon introducing geometric realizations for $\mathbf{H}$ and $\mathbf{H}_t$, it will be natural to add a finite number of small constraints to obtain a well-defined $\mathbf{C}_t$ − "the" contraction operator.

Consider now a (smooth, K-finite) vector $f \in \mathbf{H}$, and set $f_t = \mathbf{C}_t f$. Remember that the aim of these notes is to understand the relationship between $\mathbf{H}$ and the Hilbert space $\mathbf{H}_0$ which carries the irreducible representation of G_0 attached to δ.

Is it possible that as t goes to zero, f_t should have a limit f_0 in some sense, and that f_0 should belong to $\mathbf{H}_0$? It is, but since f and f_t seem to live in different spaces, we have to be careful about what the "limit" means. In the rest of these notes, we shall use (well-known) geometric realizations which make it possible to embed the smooth, K-finite vectors of each $\mathbf{H}_t$ in a fixed Fréchet space, and prove that for its Fréchet topology, f_t has a limit f_0 as t goes to zero. From the limits thus obtained we get a vector space; if we take care to renormalize the imaginary part of the infinitesimal character for the representation of G_t considered at each t, the vector space will turn out to have a natural G_0-module structure, and to be isomorphic with $\mathbf{M}_0(\delta)$.

For the convenience of the reader, let me include a general statement that summarizes a rather large part of the work to come (with the exception of sections 4.1b and 4.2b below).

Theorem 3.2. *Suppose $\delta = (\chi, \mu)$ is a nice Mackey datum. Then there is a Fréchet space $\mathbf{E}$, a finite collection of continous linear functionals $\alpha_i \in \mathbf{E}'$, there is a vector subspace $\mathbf{H}$ of E, a map $\pi : G \to End(\mathbf{E})$, and for each $t > 0$ there is a vector subspace $\mathbf{H}_t \subset E$ and there are maps $\pi_t, \Pi_t : G_t \to End(\mathbf{E})$, which have the following properties.*

1. *The vector subspace $\mathbf{H}_t$ is π_t-and Π_t-stable; $(\mathbf{H}_t, \pi_t)$ is a tempered irreducible representation of G_t with class $\mathbf{M}_t(\chi, \mu)$, while $(\mathbf{H}_t, \Pi_t)$ is a tempered irreducible representation of G_t with class $\mathbf{M}_t(\frac{\chi}{t}, \mu)$;*
2. *There is exactly one linear map from $\mathbf{E}$ to itself which sends $\mathbf{H}$ to $\mathbf{H}_t$ and restricts to a contraction map between $(\mathbf{H}, \pi)$ and $(\mathbf{H}_t, \pi_t)$, while satisfying $\alpha_i \circ \mathbf{C}_t = \alpha_i$ for all i. The family $(\mathbf{C}_t)_{t>0}$ is then weakly continuous.*
3. *For each $f \in \mathbf{E}$, there is a limit (in $\mathbf{E}$) to $\mathbf{C}_t f$ as t goes to zero.*
4. *Define $\mathbf{H}_0$ as $\left\{ \lim_{t\to 0} \mathbf{C}_t f \;|f \in \mathbf{H}\right\}$, suppose f_0 is in $\mathbf{H}_0$, consider an element f of $\mathbf{H}$ such that $\lim_{t\to 0} \mathbf{C}_t f = f_0$, and set $f_t = \mathbf{C}_t f$. Then for each g_0 in G_0, there is a limit to $\Pi_t(\alpha_t(g_0))f_t$ as t goes to zero, this limit depends only on f_0 (and g_0), and it belongs to $\mathbf{H}_0$. Call it $\pi_0(g_0)f_0$.*
5. *We thus obtain a vector subspace $\mathbf{H}_0$ of $\mathbf{E}$, and a representation π_0 of G_0 on $\mathbf{H}_0$. This representation is then unitary irreducible, and its equivalence class is $\mathbf{M}_0(\delta)$.*

"Nice" here means that if $\delta = (\chi, \mu)$, the irreducible-tempered-representation-with-real-infinitesimal-character $\mathbf{V}_{M_\chi}(\mu)$ is either a discrete series or a limit of discrete series representation of $\mathbf{M}_\chi$, or that μ is trivial. We will of course see (in section 6.2) why I have not been able to remove this restriction: for other Mackey data, I can but conjecture Theorem 3.2 or prove a less satisfactory version in which the outcome $\mathbf{H}_0$ of the contraction (point 5.) is reducible (and splits as a countable direct sum of irreducible G_0-modules), but contains $\mathbf{M}_0(\delta)$ as a multiplicity-one closed subspace.

The space $\mathbf{E}$ will roughly be a space of continuous functions with values in a finite-dimensional vector space, and the constraint enforced by the linear functionals will be that $\mathbf{C}_t$ preserve the value of functions at a distinguished point. The relationship between π_t and Π_t will be a simple renormalization of the continuous parameters (imaginary parts of the infinitesimal characters) for the representations: if the equivalence class of π_t is that of $\mathrm{Ind}_{P_{\chi,t}}^{G_t}\left(\mathbf{V}_{M_{\chi,t}}(\mu) \otimes e^{i\chi}\right)$, that of Π_t will be $\mathrm{Ind}_{P_{\chi,t}}^{G_t}\left(\mathbf{V}_{M_{\chi,t}}(\mu) \otimes e^{i\frac{\chi}{t}}\right)$.

The above announcment is meant to make it easier to follow the upcoming sections, but it is rather vague ; I shall of course include more precise statements along the way − see especially Theorem 7.2. below. In Section 4.2b, we shall also see that in some instances there are other natural ways to discuss the contraction: there the renormalization of infinitesimal characters will correspond to a geometrical procedure taking place at the level of vectors (see remark 4.3.2).

4 Principal series representations

In this section, we choose a Mackey datum $\delta = (\lambda, \mu)$ with regular λ; the representations of G with class $\mathbf{M}(\delta)$ are unitary principal series representations, and several existing results can be understood as giving flesh to Mackey's analogy at the level of carrier spaces. I will comment on some of them in section 4.3 below.

4.1 Two geometric realizations

There are several well-known function spaces carrying a representation of G with class $\mathbf{M}(\delta)$ − see for instance section VII.1. in [22]. I will use two of these function spaces here: in the first, the functions are defined on K/M − which has the same meaning in G and G_0; in the second, they are defined on G/K, or equivalently on $\mathfrak{p}$, and the geometrical setting in section 2.2 will prove helpful.

Before I proceed to the contraction, let me describe the corresponding realizations of $\mathbf{M}(\delta)$. They are famous, of course.

4.1.a The compact picture

Since I will use this realization outside the principal series, until the end of this subsection I do not assume that $\delta = (\lambda, \mu)$ has λ regular. I write $P = MAN$ for the cuspidal parabolic subgroup we induce from.

Let me write V_σ for the space of a tempered irreducible M-module of class $\sigma = \mathbf{V}_M(\mu)$, and suppose that an M-invariant inner product is fixed on V_σ. A possible Hilbert space for $\mathbf{M}(\delta)$ is

$$\mathbf{H}_\sigma^{comp} = \{f \in \mathbf{L}^2(K; V_\sigma) \mid f(km) = \sigma(m)^{-1} f(k), \forall (k, m) \in K \times M\}.$$

To say how G acts on $\mathbf{H}_\sigma^{comp}$ I need the Iwasawa projections κ, $\mathbf{m}$, $\mathbf{a}$, ν sending an element of G to the unique quadruple

$$(\kappa(g), \mathbf{m}(g), \mathbf{a}(g), \nu(g)) \in K \times \exp_G(\mathfrak{m} \cap \mathfrak{p}) \times \mathfrak{a} \times N$$

such that $g = \kappa(g)\mathbf{m}(g)\exp_G(\mathfrak{a}(g))\nu(g)$ (this quadruple is unique, see [22]). Note that if P is minimal, the map $\mathbf{m}$ is trivial. The operator for the action of $g \in G$ on $\mathbf{H}_\sigma^{comp}$ is then

$$\pi_{\lambda,\mu}^{comp}(g) = f \mapsto \left[k \mapsto \exp\langle -i\lambda - \rho, \mathbf{a}(g^{-1}k)\rangle \sigma(\mathbf{m}(g^{-1}k))^{-1} f\left(\kappa(g^{-1}k)\right)\right]$$

where ρ is the half-sum of those roots of $(\mathfrak{g}, \mathfrak{a})$ that are positive in the ordering used to define N. Note that the Hilbert space does not depend on λ, but that the G-action does.

It will be useful to recall how this is related to the usual "induced picture", for which the Hilbert space is

$$\mathbf{H}_\delta^{ind} = \left\{f : G \to V_\sigma \mid f(gme^H n) = e^{\langle -i\lambda - \rho, H\rangle}\sigma(m)^{-1} f(g) \text{ for } (g, me^H n) \in G \times P, \text{ and } f\big|_K \in \mathbf{L}^2(K; V_\sigma)\right\},$$

the inner product is the $\mathbb{L}^2$ scalar product between restrictions to K, and the G-action is given by $\pi_\delta^{ind}(g)f = [x \mapsto f(g^{-1}x)]$ for (g, f) in $G \times \mathbf{H}_\delta^{ind}$. Because an element of $\mathbf{H}_\delta^{ind}$ is completely determined by its restriction to K thanks to its P-equivariance, restriction to K induces an isometry (say $\mathbf{R}$) between $\mathbf{H}_\delta^{ind}$ and $\mathbf{H}_\sigma^{comp}$; the definition of $\pi_{\lambda,\mu}^{comp}$ is just what is needed to make $\mathbf{R}$ an intertwining operator.

4.1.b Helgason's waves and picture for the spherical principal series

Let me assume again that $\delta = (\lambda, \mu)$ has λ regular, and suppose in addition that μ is the trivial representation of M. Then there is a distinguished element in $\mathbf{H}_\sigma^{comp}$: the constant function on K with value one. Under the isometry $\mathbf{R}$, it corresponds to the function

$$\bar{e}_{\lambda,1} = ke^H n \mapsto e^{\langle -i\lambda - \rho, H\rangle}$$

in $\mathbf{H}_\delta^{ind}$, which in turn defines a function on G/K if we set $e_{\lambda,1}(gK) = \bar{e}_{\lambda,1}(g^{-1})$, and a function on $\mathfrak{p}$ if we set $e_{\lambda,1}(v) = e_{\lambda,1}(\exp_G(v)K)$.

Here is a plot of $e_{\lambda,1}$ when G is $SL_2(\mathbb{R})$:

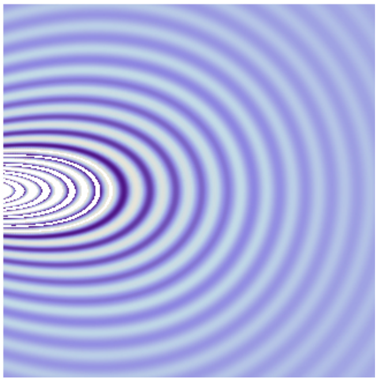


Figure 2: Plot of the real part of the Helgason wave $e_{30,1}$. I used the mapping from $\mathbb{R}^2$ to the unit disk provided by the Cartan decomposition, and the explicit formulae availiable on the unit disk: see [15], chapter 0. The x- and y- range is [-1.5, 1.5] (this region is chosen so that the modulus varies clearly but within a displayable range, and the choice of λ is to have enough waviness in the region).

Now set $e_{\lambda,b}(v) = e_{\lambda,1}(b^{-1}v)$ for b in K/M and v in $\mathfrak{p}$. Then

$$\begin{aligned} \mathbf{L}^2(K/M) = \mathbf{H}_\sigma^{comp} &\to \mathbf{C}^\infty(\mathfrak{p}) \\ F &\mapsto \int_{K/M} e_{\lambda,b} F(b) db \end{aligned}$$

turns out to intertwine $\pi_{\lambda,1}^{comp}$ with the quasi-regular action of G on $\mathbf{C}^\infty(\mathfrak{p})$, and to be an injection (see [16], Chapter 3). I shall write $\mathbf{H}_\lambda^{Helgason}$ for the image of this map; of course it inherits a Hilbert space structure from that of $\mathbf{H}_\lambda^{comp}$.

4.2 The contraction operators

4.2.a Contraction of principal series representations in the compact picture

Let me consider here the principal series representation $\pi_{\lambda,\sigma}^{t,comp}$ of G_t (here λ is regular and σ is any element of $\widehat{M_{\mathfrak{p}}}$) which acts on $\mathbf{H}_\sigma^{comp} = \mathbf{L}^2(K, V_\sigma)$. We can define a representation of G as the composition

$$\varpi_{\lambda,\sigma}^{t,comp} \;\; G \xrightarrow{\varphi_t^{-1}} G_t \xrightarrow{\pi_{\lambda,\sigma}^{t,comp}} \mathrm{End}(\mathbf{H}_\sigma^{comp}).$$

The next lemma indicates how (λ, σ) is to be interpreted as a Mackey datum for G_t :

Lemma 4.1. *For each $t > 0$, $\varpi_\lambda^{t,comp}$ is equal to $\pi_{\lambda/t}^{comp}$.*

To prove this lemma, we need only write down the details for the definition of $\pi_{\lambda/t,\sigma}^{t,comp}$. We have to understand what happens to the half-sum of positive roots when we go from G to G_t, and to make the relationship between the Iwasawa decompositions in both groups clear. Here is a first step :

Lemma 4.2. *If $\alpha \in \mathfrak{a}^\star$ is a root of $(\mathfrak{g}, \mathfrak{a})$, then $t \cdot \alpha$ is a root of $(\mathfrak{g}_t, \mathfrak{a})$.*

Proof. When α is a root of $(\mathfrak{g}, \mathfrak{a})$, there is a nonzero $X \in \mathfrak{g}$ such that $[X, H] = \alpha(H)X$ for each $H \in \mathfrak{a}$. To keep track of X through the contraction, let's write $X = X_e + X_h$ with $X_e \in \mathfrak{k}$ and $X_h \in \mathfrak{p}$. Then

$$[X_e, H] - \alpha(H)X_h = [X_h, H] - \alpha(H)X_e. \tag{4.1}$$

The left-hand-side of (4.1) is in $\mathfrak{p}$ and the right-hand-side is in $\mathfrak{k}$, so both are zero.

Now, the isomorphism ϕ_t^{-1} sends X to $X^t = X_e + \frac{1}{t}X_h \in \mathfrak{g}_t$, and for each $H \in \mathfrak{a}$,

$$[X^t, H]_{\mathfrak{g}_t} = [X_e, H]_{\mathfrak{g}_t} + \frac{1}{t}[X_h, H]_{\mathfrak{g}_t} = [X_e, H]_{\mathfrak{g}} + t[X_h, H]_{\mathfrak{g}}. \tag{4.2}$$

(for the last equality, recall that if U, V are in $\mathfrak{g}$, the bracket $[U, V]_{\mathfrak{g}_t}$ is defined as $\phi_t^{-1}[\phi_t U, \phi_t V]_{\mathfrak{g}}$, hence $[U, V]_{\mathfrak{g}_t} = [U, V]_{\mathfrak{g}}$ when U lies in $\mathfrak{k}$ and V lies in $\mathfrak{p}$, and $[U, V]_{\mathfrak{g}_t} = t^2[U, V]_{\mathfrak{g}}$ when they both lie in $\mathfrak{p}$).

The right-hand side of (4.2) is $\alpha(H)X_h + t \cdot \alpha(H)X_e = t \cdot \alpha(H)X^t$, so X^t is in the $(\mathfrak{g}_t, \mathfrak{a})$ root space for $t \cdot \alpha$, which proves lemma 4.2. □

The proof shows that the root space for $t \cdot \alpha$ is the image of $\mathfrak{g}_\alpha$ under ϕ_t^{-1}; a consequence of this is that the subgroups M_t, A_t and N_t of G_t provided by the constructions of section 3.2 are the images of M, A and N under φ_t^{-1}. If $g = k \exp_G(H)n$ is the Iwasawa decomposition of $g \in G$, the corresponding Iwasawa decomposition of $\varphi_t^{-1}g$ is then $\varphi_t^{-1}g = k \cdot \varphi_t^{-1}\left[\exp_G(H)\right] \cdot \varphi_t^{-1}(n)$. Thus

$$\kappa_t(\varphi_t^{-1}g) = \kappa(g);$$

$$\mathbf{a}_t(\varphi_t^{-1}g) = \frac{\mathbf{a}(g)}{t}.$$

The second equality uses the commutation relation between group exponentials and group morphisms. Now, because of Lemma 4.2, for each $\gamma \in G_t$ we know that

$$\pi_{\lambda,\sigma}^{t,comp}(\gamma) = f \mapsto \left[k \mapsto \exp\langle -i\lambda - t\rho, \mathbf{a}_t(\gamma^{-1}k)\rangle f\left(\kappa_t(\gamma^{-1}k)\right)\right].$$

Hence

$$\pi_{\lambda,\sigma}^{t,comp}(\varphi_t^{-1}(g)) = f \mapsto \left[k \mapsto \exp\langle -i\lambda - t\rho, \mathbf{a}_t(\left[\varphi_t^{-1}g\right]^{-1}k)\rangle f\left(\kappa_t(\left[\varphi_t^{-1}g\right]^{-1}k)\right)\right].$$

And rearranging,

$$\begin{aligned}
\pi_{\lambda,\sigma}^{t,comp}(\varphi_t^{-1}(g)) &= f \mapsto \left[k \mapsto \exp\langle -i\frac{\lambda}{t} - \rho, t \cdot \mathbf{a}_t(\varphi_t^{-1}\left[g^{-1}k\right])\rangle f\left(\kappa_t(\varphi_t^{-1}\left[g^{-1}k\right])\right)\right] \\
&= f \mapsto \left[k \mapsto \exp\langle -i\frac{\lambda}{t} - \rho, \mathbf{a}(g^{-1}k)\rangle f\left(\kappa(g^{-1}k)\right)\right] \\
&= \pi_{\frac{\lambda}{t},\sigma}^{comp}(g),
\end{aligned}$$

so the proof of lemma 4.1 is complete. □

To discuss the contraction from G to G_0 the situation seems disappointingly trivial here: the Hilbert space is the same for each t, including $t = 0$, and because of Lemma 4.1 the natural "contraction" operator $\mathbf{C}_t$ is the identity. However, this does not mean that

Mackey's analogy is devoid of interest for the principal series, even from the point of view of Hilbert spaces; the interplay with Helgason's picture will show this clearly, but let us linger in the compact picture for a moment.

I can use the diffeomorphisms $\alpha_t : G_0 \to G_t$ which realize the Cartan decomposition (see section 2.1) to define maps $\tilde{\pi}_t$ from G_0 to $\mathrm{End}(\mathbf{H}_\sigma^{comp})$, setting

$$\tilde{\pi}^t = \pi_{\lambda,\sigma}^{t,comp} \circ \alpha_t.$$

Because of lemma 4.1, $\tilde{\pi}_t(g_0)$ is an operator for a principal series representation of G with infinitesimal character $\frac{i\lambda}{t}$; but as t goes to zero it gets closer and closer to an operator for the representation of G_0 with Mackey datum (λ, σ):

Theorem 4.1. *For each g_0 in G_0, there is a limit to $\tilde{\pi}_t(g_0)$ as t goes to zero; it is the operator $\pi_0(g_0)$. The convergence holds both in the usual weak sense when the operators are viewed as unitary operators on $\mathbb{L}^2(K)$, and in the weak topology associated to that of uniform convergence on $\mathbf{C}(K)$.*

To prove this theorem, recall that

$$\pi_{\lambda,\sigma}^{t,comp}(\exp_{G_t}(v)k) = f \mapsto \left[u \mapsto \exp\left\langle -i\lambda - t\rho, \mathbf{a}_t\left((k^{-1} \cdot_t \exp_{G_t}(-v) \cdot_t u)\right)\right\rangle f\left(\kappa_t(k^{-1}\exp_{G_t}(-v)u)\right)\right].$$

Here the products are products in G_t. On the other hand, recall from section 2.2 that

$$\pi_0(k,v) = f \mapsto \left[u \mapsto \exp\left\langle i\lambda, Ad(u^{-1})v\right\rangle f\left(k^{-1}u\right)\right].$$

To make the two look more similar, notice that

$$\begin{aligned}\left[\pi_{\lambda,\sigma}^{t,comp}(\exp_{G_t}(v)k)f\right](u) &= e^{\langle -i\lambda - t\rho, \mathbf{a}_t\left((k^{-1}u)\cdot_t \exp_{G_t}(-Ad(u^{-1})v)\right)\rangle} f\left(\kappa_t(k^{-1}u\exp_{G_t}(-Ad(u^{1})v)\right)\\ &= e^{\langle -i\lambda - t\rho, \mathbf{a}_t\left(\exp_{G_t}(-Ad(u^{-1})v)\right)\rangle} f\left((k^{-1}u)\cdot\kappa_t(exp_{G_t}(-Ad(u^{-1})v))\right).\end{aligned}$$

We now need to see how the Iwasawa projection parts behave as t goes to zero. Let us write $\mathfrak{K}$ and $\mathfrak{I}$ for the maps from $\mathfrak{p}$ to $\mathfrak{a}$ sending $v \in \mathfrak{p}$ to the K- and $\mathfrak{a}$-Iwasawa components of $\exp_G(v)$, respectively (so $\mathfrak{K}$ is $\kappa \circ \exp_G$ and $\mathfrak{I}$ is $\mathbf{a} \circ \exp_G$); let us likewise set $\mathfrak{K}_t = \kappa_t \circ \exp_{G_t}$ and $\mathfrak{I}_t = \mathbf{a}_t \circ \exp_{G_t}$. The Iwasawa map $\mathfrak{I}_t$ from $\mathfrak{p}$ to $\mathfrak{a}$ is a nonlinear map, but as t goes to zero it gets closer and closer to a linear projection :

Lemma 4.3. *As t tends to zero, $\mathfrak{I}_t$ admits as a limit (in the sense of uniform convergence on compact subsets of $\mathfrak{p}$) the orthogonal projection from $\mathfrak{p}$ to $\mathfrak{a}$, while $\mathfrak{K}_t$ tends to the constant function on $\mathfrak{p}$ with value $\mathbf{1}_K$.*

Proof. We will check now that $\mathfrak{I}_t$ is none other than $v \mapsto \frac{1}{t}\mathfrak{I}(tv)$. Since φ_t is a group morphism from G_t to G, the definition of group exponentials does imply that $\exp_G(tv) = \exp_G(d\varphi_t(1)v) = \varphi_t\left(\exp_{G_t} v\right)$. Let us write $\exp_{G_t} v = ke^{\mathfrak{I}_t(v)}n_t$ with $k \in K$ and $n_t \in N_t$, then $\varphi_t\left(\exp_{G_t} v\right) = ke^{t\mathfrak{I}_t(v)}n$, with $n = \varphi_t(n_t)$ in N. So we know that

$$\exp_G(tv) = ke^{t\mathfrak{I}_t(v)}n \tag{4.3}$$

and thus that $\mathfrak{I}(tv) = t\mathfrak{I}_t(v)$, as announced.

But then as t goes to zero, the limit of $\mathfrak{I}_t(v)$ is the value at v of the derivative $d\mathfrak{I}(0)$. Now this does yield the orthogonal projection of v on $\mathfrak{a}$: although the Iwasawa decomposition of $\mathfrak{g}$ is not an orthogonal direct sum because $\mathfrak{k}$ and $\mathfrak{n}$ are not orthogonal to each other, they are both orthogonal to $\mathfrak{a}$ with respect to the Killing form of $\mathfrak{g}$, so the direct sum $\mathfrak{k} \oplus \mathfrak{n}$ is the orthogonal of $\mathfrak{a}$.

As for $\mathfrak{K}_t$, from (4.3) we see that $\mathfrak{K}_t(v) = \kappa_t\left(exp_{G_t} v\right) = \kappa\left(\exp_G(tv)\right)$, and this does go to the identity uniformly on compact subsets t goes to zero (here I measure distances on K with the bi-invariant metric on K whose volume form is the normalized Haar measure). $\square$

The "weak convergence with respect to the topology of uniform convergence" part of Theorem 4.1 follows immediately, and since we are dealing with continuous functions on a compact manifold here, uniform convergence implies $\mathbb{L}^2$ convergence. This concludes the proof of Theorem 4.1; for remarks, see 4.3.1 below. $\square$

4.2.b Contraction of the spherical principal series in Helgason's picture

For each $t > 0$, each $\lambda \in \mathfrak{a}^\star$ and $b \in K/M$, define a function on $\mathfrak{p}$ by setting

$$\varepsilon^t_{\lambda,b}(v) = e^{\langle\ i\lambda + t\rho\ ,\ \mathfrak{I}_t(Ad(b)\cdot v)\ \rangle} \text{ for } v \text{ in } \mathfrak{p}.$$

Let me simplify the notations a bit and write $\mathbf{H}^t_\lambda$ for the Hilbert space $\mathbf{H}^{t,Helgason}_\lambda$ which one can associate to G_t as in section 4.1. Let me also set $B = K/M$. The next lemma gathers some simple consequences of Lemma 4.1, Lemma 4.3 and their proofs.

Lemma 4.4. *1. The Hilbert space $\mathbf{H}^t_\lambda$ is exactly $\left\{\int_B \varepsilon^t_{\lambda,b} F(b)db \mid F \in \mathbb{L}^2(B)\right\}$.*

2. For each $\lambda \in \mathfrak{a}^\star$ and each $b \in K/M$, the Helgason waves $\varepsilon^t_{\lambda,b}$ converge uniformly on compact subsets of $\mathfrak{p}$ to the Euclidean plane wave $v \mapsto \exp\left(\langle i\lambda, Ad(b)\cdot v\rangle\right)$ as t goes to zero.

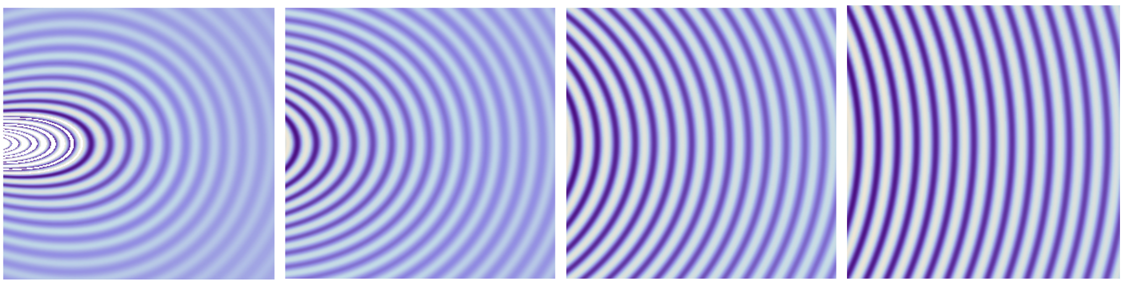


Figure 3: Illustration of lemma 4.4: these are plots of $\varepsilon^{1/2^k}_{\lambda,1}$, $k = 0, 1, 2, 3$, in the same domain as in Figure 2. Each of these waves is a building block for a principal series representation of G whose continuous parameter is $2^k\lambda$, with $\lambda = 30$ here.

Lemma 4.4. is not an unpleasant way to understand how the principal series representation $\mathbf{M}(\lambda, 1)$ is related with the representation $\mathbf{M}_0(\lambda, 1)$ of G_0, using the contraction from G_t to G_0. In addition, the above reformulation of Helgason's realization as a space of functions on $\mathfrak{p}$ makes it possible to understand the contraction process in more geometrical terms, in the spirit of what we are going to do for the discrete series below.

Recall that G_t acts on $\mathbf{H}^t_\lambda$ via $\pi_t(\gamma) = (f \mapsto [x \mapsto f(\gamma^{-1} \cdot_t x)])$ for γ in G_t. Define $\varpi_{t,\lambda}$ as $G \to \mathrm{End}(\mathbf{H}^t_\lambda)$ as $\pi_t \circ \varphi_t^{-1}$, and

$$\mathbf{C}^\lambda_t : \mathbf{H}_{\lambda/t} \to \mathbf{H}^t_\lambda$$
$$\int_B e_{\frac{\lambda}{t},b} F(b)db \mapsto \int_B \varepsilon^t_{\lambda,b} F(b)db.$$

Then $\mathbf{C}^\lambda_t$ does intertwine $\pi_{\lambda/t}$ and $\varpi_{t,\lambda}$: to see this, notice that

$$\varepsilon^t_{\lambda,b} = v \in \mathfrak{p} \mapsto e^{\langle\, i\frac{\lambda}{t}+\rho \ ,\ \Im(b\cdot(tv))\,\rangle} = \varepsilon^1_{\frac{\lambda}{t},b}(tv) = \varepsilon^1_{\frac{\lambda}{t},b}(z_t^{-1}v) = e_{\frac{\lambda}{t},b}(z_t^{-1}v)$$

so when f is an element of $\mathbf{H}_{\lambda/t}$, $\mathbf{C}^\lambda_t f$ is none other than $f \circ z_t^{-1}$. Using the fact that $\varphi_t^{-1}(g) \cdot z_t(v)$ is equal to $z_t(g \cdot v)$ for all v, we see that for every g in G,

$$\varpi_{t,\lambda}(g)\left(e_{\frac{\lambda}{t},b} \circ z_t^{-1}\right) = \left(\pi_{\lambda/t}(g) e_{\frac{\lambda}{t},b}\right) \circ z_t^{-1},$$

that is,

$$\varpi_{t,\lambda}(g)\left(\mathbf{C}^\lambda_t e_{\frac{\lambda}{t},b}\right) = \mathbf{C}^\lambda_t \left[\pi_{\lambda/t}(g) e_{\frac{\lambda}{t},b}\right].$$

Note that strictly speaking $\varepsilon^t_{\lambda,b}$ is not in $\mathbf{H}^t_\lambda$, but the definition of $\mathbf{C}^\lambda_t$ can be extended to $\left\{\int_B \varepsilon^t_{\lambda,b} F(b)db \mid F \text{ is a distribution on } B\right\}$.

Because every element of $\mathbf{H}_{\lambda/t}$ is a combination of the $e_{\frac{\lambda}{t},b}$ and the G-action commutes with the way the combinations are built, this does yield

$$\varpi_{t,\lambda}(g)(\mathbf{C}^\lambda_t f) = \mathbf{C}^\lambda_t(\pi_{\lambda/t}(g)f)$$

as announced.

Our contraction operator $\mathbf{C}^\lambda_t$ is the only intertwining operator between $\pi_{\lambda/t}$ and $\varpi_{t,\lambda}$ which preserves the linear functional isolating the value of functions at zero. Because of lemma 4.4 (ii), we see that $\mathbf{C}^\lambda_t\left(\int_B e_{\frac{\lambda}{t},b} F(b)db\right) = \int_B \varepsilon^t_{\lambda,b} F(b)db$ converges to $\int_B e^{\langle i\lambda, Ad(b)\cdot v\rangle} F(b)db$, a square-integrable, smooth function on $\mathfrak{p}$ whose Fourier transform is concentrated on $\mathrm{Ad}^\star(K)\cdot\chi$.

As I recalled in section 2.3, the vector space

$$\mathbf{H}_0 := \left\{ v \mapsto \int_B e^{\langle i\lambda, Ad(b)\cdot v\rangle} F(b)db \mid F \in \mathbf{L}^2(B)\right\}$$

with the G_0-action inherited from that of G_0 on $\mathfrak{p}$, is an irreducible G_0-module with class $\mathbf{M}_0(\lambda, 1)$.

Let us write $\mathbf{R}_t : \mathbf{H}^t_{t\lambda} \to \mathbf{H}^t_\lambda$ for the map which sends $\int_B e_{\frac{t\lambda}{t},b} F(b)db$ to $\int_B e_{\frac{\lambda}{t},b} F(b)db$. The composition $\mathbf{R}_t \circ \mathbf{C}^{t\lambda}_t$ is a map from $\mathbf{H}_\lambda$ to $\mathbf{H}^t_\lambda$, though it is not an intertwining operator between π_λ and $\varpi_{t,\lambda}$. Since we saw that $\mathbf{C}^\lambda_t$ does not depend on λ, we can rewrite the composition $\mathbf{R}_t \circ \mathbf{C}^{t\lambda}_t$ as $\mathbf{R}_t \circ \mathbf{C}_t$: this is an operator from $\mathbf{H}_\lambda$ to $\mathbf{H}^t_\lambda$ which zooms in on the value of functions close to zero and at the same time renormalizes the frequencies in Helgason's waves. We can then summarize the above discussion with the above statement.

Theorem 4.2. *For each $f \in \mathbf{H}_\lambda$, there is a limit to $(\mathbf{R}_t \circ \mathbf{C}_t)f$ for the topology of uniform convergence on compact subsets of $\mathfrak{p}$, and this limit belongs to $\mathbf{H}_0$. In fact, $f \mapsto \lim\limits_{t\to 0}(\mathbf{R}_t \circ \mathbf{C}_t)f$ defines a linear, K-invariant isometry between $\mathbf{H}_\lambda$ and $\mathbf{H}_0$.*

4.3 Three remarks

4.3.1. Theorem 4.1 can be viewed as a reformulation of Theorem 1 in Dooley and Rice's paper [8]. If I include it to these notes it is because I think the interplay with Helgason's picture throws some light on the phenomenon, because section 7 below will be a simple-but-technical adaptation of the strategy in section 4.2.1, and because all the ingredients in the proof of Lemma 4.1 and Theorem 4.1 will serve again in section 7.

4.3.2. The discussion in Section 4.2.b shows that, given a Mackey datum δ, there are several possible settings to discuss the contraction from a representation of G with class $\mathbf{M}(\delta)$ to one of G_0 with class $\mathbf{M}_0(\delta)$ $\mathbf{E}$; Theorem 3.2, in which the renormalization of infinitesimal characters is introduced in the G_t-actions rather than the Hilbert spaces, is but one of them. Let me give some precisions on the way Helgason's picture provides a setting that is not unlike that of Theorem 3.2, save for the fact that the naturality of our contracting operators is cannot be justified here by their being uniquely determined through the "contraction map" requirement.

Let $\mathbf{E}$ be the Fréchet space of smoth functions on $\mathfrak{p}$. Then $\mathbf{H}^t_\lambda$ is a vector subspace of $\mathbf{E}$ for each $t > 0$ and each λ; in addition, there is a measure μ on $\mathfrak{a}^\star$ with the property that every smooth and compactly supported function f on $\mathfrak{p}$ can be written as $\int_{\mathfrak{a}^\star} f_\lambda d\mu(\lambda)$, with f_λ in $\mathbf{H}_\lambda$.

We defined a contraction map $\mathbf{C}^{t\lambda}_t$ from $\mathbf{H}_\lambda$ to $\mathbf{E}$ in section 4.2.2, so we can define a linear operator from $\mathcal{C}^\infty_c(\mathfrak{p})$ to $\mathbf{E}$ by setting

$$\begin{aligned} &\mathbf{\Gamma}_t : \mathcal{C}^\infty_c(\mathfrak{p}) \to \mathbf{E} \\ &\int_{\mathfrak{a}^\star} f_\lambda d\mu(\lambda) \mapsto \int_{\mathfrak{a}^\star} \left(\mathbf{R}_t \circ \mathbf{C}^{t\lambda}_t\right) f_\lambda d\mu(\lambda), \end{aligned}$$

and then use the density of $\mathcal{C}^\infty_c(\mathfrak{p})$ in $\mathbf{E}$ to extend it to a map, still denoted Γ_t, from $\mathbf{E}$ to $\mathbf{E}$ (it is a consequence of the results of the previous section that the extension does work). Although the restriction of $\int_{\mathfrak{a}^\star} f_\lambda d\mu(\lambda) \mapsto \int_{\mathfrak{a}^\star} \mathbf{C}_t f_\lambda d\mu(\lambda)$ to $\mathbf{H}_\lambda$ is uniquely determined by the requirement that it be a contraction map which preserves the value of functions at zero, the renormalization of frequencies means that $\mathbf{\Gamma}_t$ does *not* intertwine the G-and G_t-actions. But the discussion leading up to Theorem 4.2 does show that for each f in $\mathbf{E}$, there is a limit in $\mathbf{E}$ to $\mathbf{\Gamma}_t f$ as t goes to zero, and that starting with an element of $\mathbf{H}_\lambda$ produces in the limit an element of $\mathbf{H}_0$. Because of Lemma 2.3, the action of G_t on $\mathbf{H}^t_\lambda$ does in the limit yield the appropriate action of G_0 on $\mathbf{H}_0$.

4.3.3. Many existing studies compare harmonic analysis for functions on G/K with ordinary Fourier analysis for functions on $\mathfrak{p}$, with the hope of solving some apparently difficult problems on G/K, like the existence of fundamental solutions for G-invariant partial differential equations on G/K: see for instance [10, 20, 16], and especially Rouvière's book [35]. The contents of this section provide a way to turn a function on G/K into a function on $\mathfrak{p}$ in a relatively natural manner which uses the fine structure of G. Is it possible that this transformation should be related to some of the issues in [35] ? I have not looked closely into the matter at present.

5 The discrete series

In this section G will be connected, semisimple, with finite center, and I will assume that G and K have equal ranks, so that G has a nonempty discrete series. Let me again write T for a maximal torus in K.

5.1 Square-integrable solutions of the Dirac equation

Let us start with a class $\mu \in \widehat{K}$. If the highest weight of μ lies sufficiently far away from the root hyperplanes (I will make this precise immediately), the representation $\mathbf{V}_G(\mu)$ belongs to the discrete series of G. In this subsection I recall some results of Parthasarathy, Atiyah and Schmid [1, 32] which provide a Hilbert space for $\mathbf{V}_G(\mu)$.

I use standard terminology here and say that an element of $i\mathfrak{t}^\star$ is in Λ if it is the derivative of a character of T; let me write Δ_c for the set of roots of $(\mathfrak{k}_\mathbf{C}, \mathfrak{t}_\mathbf{C})$, and Δ for the set of roots of $(\mathfrak{g}_\mathbf{C}, \mathfrak{t}_\mathbf{C})$; of course $\Delta_c \subset \Delta$.

If P is any system of positive roots for Δ, we can consider the half-sum ρ_P of positive roots, and then set $\Lambda_\rho = \Lambda + \rho_P$; it is a translate of a lattice in $i\mathfrak{t}^\star$, which does not depend on which positive system P we chose in Δ.

Suppose a positive root system P has been chosen, and form the corresponding Harish-Chandra parameter $\vec{\lambda}$ of $\mathbf{V}_G(\mu)$; it is an element of $\mathfrak{t}^\star$. Write ρ_c for the half-sum of positive, compact roots, and ρ_n is the half-sum of positive, noncompact roots, so that $\rho_c + \rho_n = \rho_P$. The highest weight $\vec{\mu}$ of μ associated with the given ordering of $\mathfrak{t}^\star$ is then related with $\vec{\lambda}$ through $\vec{\lambda} = \vec{\mu} + \rho_n - \rho_c$; the fact that $\mathbf{V}_G(\mu)$ lies in the discrete series of G then translates as

$$\vec{\lambda} \text{ is a regular element of } i\mathfrak{t}^\star \tag{5.1}$$

(see the proof of lemma 3.1 for the definition).

Recall that the condition that G and K have equal ranks guarantees that $\dim(G/K)$ is an even integer, say $2q$. Let us write S for a 2^q-dimensional space on which $Spin(2q)$ acts through the spinor representation. The module S splits into two irreducible 2^{q-1}-dimensional $Spin(2q)$-submodules S^+ and S^-, with ρ_n a weight of S^+ (for the action induced by the natural map from $\mathfrak{k}$ to $Spin(2q)$).

Suppose $V_{\mu^\flat}$ is the carrier space of an irreducible $\mathfrak{k}_\mathbb{C}$-module with highest weight $\mu^\flat = \vec{\mu} - \rho_n$. Then we can consider the tensor product $V_{\mu^\flat} \otimes S^\pm$, and although neither $V_{\mu^\flat}$ nor $S^\pm$ need be a K-module if G is not simply connected, it turns out that the action of $\mathfrak{k}$ on $V_{\mu^\flat} \otimes S^\pm$ does lift to K − the half-integral ρ_n-shifts in the weights do compensate. So we can consider the equivariant bundle $\mathfrak{E} = G \otimes_K (V_\mu \otimes S)$ over G/K, as well as equivariant bundles $\mathfrak{E}^\pm = G \otimes_K (V_\mu \otimes S^\pm)$.

Now, the natural G-invariant metric that G/K inherits from the Killing form of $\mathfrak{g}$ and the built-in G-invariant spin structure of $\mathfrak{E}$ make it possible to define a first-order differential operator D acting on smooth sections of $\mathfrak{E}$, the Dirac operator: since I will need a few immediate consequences of its definition the next subsection, let me give a quick definition, referring to [32] for details.

Suppose $(X_i)_{i=1..2q}$ is an orthonormal basis of $\mathfrak{p}$. Recall that the definition of spinors comes with a map $\mathbf{c}$ from $\mathfrak{p}_\mathbb{C}$ to $\mathrm{End}(S)$, called Clifford multiplication, such that $\mathbf{c}(X)$ sends $S^\pm$ to $S^\mp$, and that every X in $\mathfrak{p}$ defines a left-invariant vector field on G/K, which yields a first-order differential operator $X^\mathfrak{E}$ acting (componentwise in the natural trivialization associated to the action of G on G/K) on sections of $\mathfrak{E}$. The Dirac operator is then defined by

$$Ds = \sum_{i=1}^{2q} \mathbf{c}(X_i) X_i^\mathfrak{E} s$$

when s is a section of $\mathfrak{E}$. It splits as $D = D^+ + D^-$, with $D^\pm$ sending sections of $\mathfrak{E}^\pm$ to sections of $\mathfrak{E}^\mp$.

Let me now write $\mathbf{H}_\mu$ for the space of smooth, square integrable sections of $\mathfrak{E}$ which are in the kernel of D. Since D is an essentially self-adjoint elliptic operator, $\mathbf{H}_\mu$ is a closed subspace of the Hilbert space of square-integrable sections of $\mathfrak{E}$. And as D is G-invariant, $\mathbf{H}_\mu$ is invariant under the natural action of G on sections of E.

Theorem (Parthasarathy, Atiyah & Schmid). *If μ satisfies the hypothesis* (5.1)*, then* $\mathbf{H}$ *carries an irreducible unitary representation of G, whose equivalence class is* $\mathbf{V}_G(\mu)$*.*

But here something happens to which we must pay very special attention: the details in Atiyah and Schmid's proof show that solutions to the Dirac equation do not explore the whole fibers, but that they are actually sections of a sub-bundle whose fiber, a K-module, is irreducible and of class μ. In clearer words, let W denote the isotypical K-submodule of $V_{\mu^\flat} \otimes S^+$ for the highest weight $\vec{\mu} = \mu^\flat + \rho_n$; the K-module W turns out to be irreducible. Let us write p_W for the isotypical (orthogonal) projection to $V_{\mu^\flat} \otimes S^+$ to W. Let $\mathbf{W}$ denote the equivariant bundle on G/K associated to W.

Proposition (Atiyah & Schmid). If a section of $\mathfrak{E}$ is a square-integrable solution of the Dirac equation, then it is in fact a section of $\mathbf{W}$.

Although this is not isolated as a proposition in Atiyah and Schmid's paper [1], it is proved and stated there very clearly; the statement contains a commentary which is quite interesting in the context of the present notes. [7]

> *We should remark that the arguments leading up to [the fact that the cokernel of the Dirac operator is zero] are really curvature estimates, in algebraic disguise. The curvature properties of the bundles and of the manifold G/K force all square-integrable, harmonic spinors to take values in a certain sub-bundle of $\mathbf{V}_\mu \otimes \mathbf{S}^+$, namely the one that corresponds to the K-submodule of highest weight $\mu + \rho_n$ in $V_\mu \otimes S^+$.*

To be complete, I should mention here that the context of the above quotation is one in which another nondegeneracy condition is imposed on μ besides that which guarantees that is is the lowest K-type of a discrete series representation. Atiyah and Schmid's arguments to remove this nondegeneracy condition in their main theorem do imply also that the above remark holds without the provisio.

5.2 Contraction of a discrete series representation to its minimal K-type

It is time to set up the stage for the contraction of a discrete series representation (I'm afraid the notation has to be a bit pedantic here if I want to reduce the hand-waving to a minimum).

Recall that in section 2.1, we used a diffeomorphism u_t between $\mathfrak{p}$ and G_t/K to make $\mathfrak{p}$ into a G_t-homogeneous space equipped with a metric η_t. We can then use the representation of K on $V_\mu \otimes S$ to build a G_t-invariant spinor bundle $\mathfrak{E}_t$ over G_t/K, use u_t to turn it into a bundle over $\mathfrak{p}$, and use the action of G_t to make this bundle trivial: this yields a bundle map, say T_t, from the bundle $u_t^\star \mathfrak{E}_t$ over $\mathfrak{p}$ to the trivial bundle $\mathfrak{p} \times (V_{\mu^\flat} \otimes S)$.

The definition of the Dirac operator makes sense for the homogeneous bundle $u_t^\star \mathfrak{E}$ over the Riemannian space $(\mathfrak{p}, \eta_t)$; once we trivialize using T_t we end up with a Dirac operator D_t', acting on $\mathbf{C}^\infty(\mathfrak{p}, V_{\mu^\flat} \otimes S)$ – and which is pushed forward by T_t-then-u_t to a constant multiple of the Dirac operator on G_t/K defined in the previous subsection. Motivated by

7. Atiyah and Schmid's μ is our $\mu^\flat$, their $\mathbf{V}_\mu \otimes \mathbf{S}^+$ is our $\mathfrak{E}$.

the end of the previous subsection, we build from D'_t an operator which acts on $\mathbf{C}^\infty(\mathfrak{p}, W)$, setting

$$\Delta_t := P_W \circ (D'_t)^2|_{\mathbf{C}^\infty(\mathfrak{p},W)}$$

where P_W is the orthogonal projection from $V_\mu \otimes S$ onto W.

Note that I need not assume that t is nonzero here: we get a G_0-invariant operator Δ_0 on the Euclidean space $(\mathfrak{p}, \eta_0)$, as well as G_t-invariant operators $\Delta_t, t > 0$, on the negatively-curved spaces $(\mathfrak{p}, \eta_t)$. I introduced the clumsy notation in order to spell out the proof of the following simple fact.

Lemma 5.1. *For each $f \in \mathbf{C}^\infty(\mathfrak{p}, W)$, the family $(\Delta_t f)_{t\geq 0}$ is continuous with respect to the topology of uniform convergence on compact sets of $\mathfrak{p}$.*

Proof. The Dirac operator is a first-order differential operator, so if I introduce $2q$ cartesian coordinates on $\mathfrak{p}$ using a linear basis, $D'_t|_{\mathbf{C}^\infty(\mathfrak{p},W)}$ will read

$$D'_t|_{\mathbf{C}^\infty(\mathfrak{p},W)} = \sum_{i=1}^{2q} A^i_t \partial_i + K_t$$

where the A^i_t, $i = 1...2q$, as well as K_t, are continuous functions from $\mathfrak{p}$ to $\mathrm{Hom}(W, V_{\mu^\flat} \otimes S)$. I now claim that it is clear from the details given on the definition of the Dirac operator acting on sections on $\mathfrak{E}$, and from the properties of u_t and T_t, that the maps $(x,t) \mapsto A^i_t(x)$ and $(x,t) \mapsto K_t(x)$ are continuous $\mathrm{Hom}(W, V_{\mu^\flat} \otimes S)$-valued maps [8] on $\mathfrak{p} \times \mathbb{R}$, which proves lemma 5.1. $\square$

Now let us start bringing G_0 into the picture. We know from the previous subsection (and from the fact that D'_t and its square have the same $\mathbb{L}^2$ kernel) that the $\mathbf{L}^2$ kernel of each Δ_t, $t > 0$, carries a discrete series representation of G_t whose minimal K-type is μ. On the other hand, the $\mathbf{L}^2$ kernel of Δ_0 is zero ! To recover the representation of G_0 which we are interested in, we should consider an extended kernel in which the constants are allowed.

Definition. *For each $t \geq 0$, the extended kernel of Δ_t is*

$$\mathbf{H}_t = \left\{ f \in \mathbf{C}^\infty(\mathfrak{p}, W) \mid \Delta_t f = 0, \text{ and there is a constant } c \in W \text{ such that } f + c \in \mathbf{L}^2(\eta_t, W) \right\}.$$

Note that when f is in $\mathbf{H}_t$, there can be only one constant c such that $f + c$ is square-integrable.

Lemma 5.2. *For $t \neq 0$, the extended kernel $\mathbf{H}_t$ coincides with the $\mathbf{L}^2$ kernel of Δ_t, whereas $\mathbf{H}_0$ is the space of constant W-valued functions on $\mathfrak{p}$.*

Proof. Let us come back to G/K and the Dirac operator D defined in subsection 5.1. Because of Parthasarathy's formula for its square, we know that there is a scalar σ such that

$$D^2 := D^- D^+ = -\Omega + \sigma$$

with Ω the Casimir operator acting on sections of $\mathfrak{E}$.

8. This is simply because the family $(T_t \circ u_t)^\star X_i^{\mathfrak{E}_t}$ of vector fields on $\mathfrak{p}$ is continuous with respect to the topology of uniform convergence on compact subsets.

Suppose a G-invariant trivialization of $\mathfrak{E}$ is chosen, so that D^2 is viewed as acting on functions from G/K to $V_{\mu^\flat} \otimes S$, and suppose $D^2 g = 0$, with $g = f + C$, $f \in \mathbf{L}^2(G/K, V_{\mu^\flat} \otimes S)$ and C a constant in $V_{\mu^\flat} \otimes S$. Then

$$\Omega f = \sigma f + \sigma C. \tag{5.2}$$

I claim that this cannot happen when C is nonzero. To see this, I use Helgason's Fourier transform for functions on G/K (see [16]). The Fourier transform of a smooth function with compact support on G/K is the function $(\lambda, b) \mapsto \int_{G/K} f(x) e_{\lambda,b}(x) dx$ on $\mathfrak{a}^\star \times K/M$, and this extends to an isometry $\mathbf{F}$ between $\mathbf{L}^2(G/K)$ and $\mathbf{L}^2(\mathfrak{a}^\star \times K/M)$ for a suitable measure on $\mathfrak{a}^\star \times K/M$. In addition, there is a notion of tempered distributions on G/K and $\mathfrak{a}^\star \times K/M$, and when f is a smooth, square-integrable function both f and Ωf are tempered distributions. Using this, the equality $\Omega f = \sigma f + \sigma C$ becomes an equality of tempered distributions on $\mathfrak{a}^\star \times K/M$, namely

$$\mathbf{F}(\Omega f) - \sigma \mathbf{F}(f) = \sigma C \delta_{(0,1M)}$$

with $\delta_{(0,1M)}$ the Dirac distribution at the point $(0, 1M)$. Of course, there are convenient transformation properties of $\mathbf{F}$ with respect to the G-invariant differential operators, and $\mathbf{F}(\Omega f)$ is actually the product of $\mathbf{F}(f)$ – an element of $\mathbf{L}^2(\mathfrak{a}^\star \times K/M)$ with a smooth function on $\mathfrak{a}^\star$. So if f were a smooth, square-integrable solution of (5.2), $\sigma C \delta_0$ would be the product of an element in $\mathbf{L}^2(\mathfrak{a}^\star \times K/M)$ with a smooth function on the same space. This can only happen if C is zero, and obviously Lemma 5.2 follows. □

As a result of lemma 5.2, each $\mathbf{H}_t$ carries an irreducible representation of G_t with minimal K-type μ, and $\mathbf{H}_0$ carries an irreducible representation of G_0 with the "right" equivalence class according to Mackey's analogy.

Let us now follow a vector through the contraction. We can of course use the G-action on $\mathfrak{p}$ to define a G-invariant Dirac operator on $C^\infty(\mathfrak{p}, W)$, and consider the Hilbert space $\mathbf{H}$ of smooth, $\mathbf{L}^2(\eta)$ solutions of the corresponding Dirac equation. Recall from section 3.4 that we are looking for a contraction operator $\mathbf{C}_t$ from $\mathbf{H}$ to $\mathbf{H}_t$.

Definition. *The* natural contraction $\mathbf{C}_t : \mathbf{H} \to \mathbf{H}_t$ *is the only contraction map (Definition 3.3) such that for each $f \in \mathbf{H}$, $(\mathbf{C}_t f)(0) = f(0)$.*

Now, we set up the geometrical stage in a way which makes it very easy to identify $\mathbf{C}_t$. Recall from lemma 2.2 that the dilation

$$z_t \; : \; x \mapsto \frac{x}{t}$$

intertwines the actions of G and G_t on $\mathfrak{p}$. As a consequence, $z_t^\star \eta_t$ is a G-invariant metric on $\mathfrak{p}$; but there are not many such metrics: since the derivative of z_t is multiplication by t^{-1} and η_1 and η_t coincide at zero, we deduce that

$$z_t^\star \eta_t = t^{-2} \cdot \eta_1 \tag{5.3}$$

(note the coherence with the fact that η_t has curvature $-t^2$, while η_1 has curvature -1).

We can use z_t to transform functions on $\mathfrak{p}$, setting

$$\mathbf{Z}_t f := x \mapsto f(t \cdot x).$$

As an immediate consequence of (5.3) and the definition of the Dirac operator, we get

$$\mathbf{Z}_t^{-1} \Delta_t \mathbf{Z}_t = t^4 \cdot \Delta_1.$$

Together with the fact that $\mathbf{Z}_t f$ is square-integrable with respect to η_t as soon as f is square-integrable with respect to η_1, this means that $\mathbf{Z}_t$ sends $\mathbf{H}_1$ to $\mathbf{H}_t$. Thus $\mathbf{Z}_t$ satisfies the properties in Definition 5.2.

So $\mathbf{C}_t$ is something very simple indeed :

Lemma 5.3. *The natural contraction $\mathbf{C}_t$ is none other than the restriction to $\mathbf{H}_t$ of the zooming-in operator $\mathbf{Z}_t$.*

Now we can get back to the program of section 3.3 and follow it to its end. Let's start with an element f of $\mathbf{H}$, and set $f_t = \mathbf{C}_t f$. We know that $\mathbf{H}$ splits as a direct sum according to K-types, in other words, we can write f as a Fourier series

$$f = \sum_{\lambda \in \widehat{K}} f_\lambda$$

where f_λ belongs to a closed subspace $\mathbf{H}^\lambda$ of $\mathbf{H}$ on which $\pi|_K$ restricts as a direct sum of copies of λ; from Harish-Chandra we know that each $\mathbf{H}^\lambda$ is finite-dimensional (and from Blattner, whose conjecture was proved by Hecht and Schmid [13], we know that there is an explicit-but-computer-unfriendly formula for its dimension).

Of course a parallel decomposition holds for $\mathbf{H}_t$, $t \neq 0$, and f_t, too, has a Fourier series

$$f_t = \sum_{\lambda \in \widehat{K}} f_{t,\lambda}.$$

Naturally the dimension of $\mathbf{H}^t_\lambda$ is independent of t, and the support of the above Fourier series does not depend on t.

The geometrical realization we chose is once more quite convenient here, because we can go a small step further and deal with each Fourier component separately :

Lemma 5.4. *The Fourier component $f_{t,\lambda}$ is actually $\mathbf{C}_t f_\lambda$.*

To prove this lemma, we need only notice that for each $\lambda \in \widehat{K}$ the map

$$f \mapsto P_\lambda f := \left[x \mapsto \int_K \xi^\star_\lambda(k)\ \mu(k) \cdot f\left(k^{-1} \cdot x\right) dk \right]$$

has a meaning as a linear operator from $\mathbf{C}^\infty(\mathfrak{p}, W)$ to itself. In the above formula ξ_λ is the global character of λ − a continuous function from K to $\mathbb{C}$ − and the star is complex conjugation. Now, if f is an element of $\mathbf{H}_t$, we can view K as a subgroup of G_t and since the adjoint action of K on $\mathfrak{p}$ is the same as that inherited from the action of G_t, the formula for $P_\lambda f$ turns out to be exactly the formula for the isotypical projection from $\mathbf{H}_t$ to $\mathbf{H}^\lambda_t$. Now we know $\mathbf{C}_t$ from lemma 5.3, and according to it P_λ obviously commutes with $\mathbf{C}_t$. This proves lemma 5.4. □

Now, on each compact subset of $\mathfrak{p}$, we know from lemma 5.3 that for the topology of uniform convergence on compact subsets of $\mathfrak{p}$, $\mathbf{C}_t f$ goes to $f(0)$ as t goes to zero. Lemma 5.4 adds the precision that each $\mathbf{C}_t f_\lambda$, $\lambda \in \widehat{K}$, goes to $f_\lambda(0)$.

Lemma 5.5. *If $\lambda \in \widehat{K}$ is different from the minimal K-type μ, then $f_\lambda(0) = 0$.*

Proof. The origin of $\mathfrak{p}$ is a fixed point for the action of K on $\mathfrak{p}$; so

$$f_\lambda(0) = (P_\lambda f)(0) = \int_K \xi^\star_\lambda(k)\ \mu(k) \cdot f(0) dk. \tag{5.4}$$

Recall that $f(0)$ is in W, which is an irreducible K-module of class μ: now, (5.4) is the formula for the orthogonal projection of $f(0)$ onto the isotypical component of W corresponding to $\lambda \in \widehat{K}$, and this projection is zero whenever $\lambda \neq \mu$. □

So each Fourier component of f, except that which corresponds to the minimal K-type, goes to zero as the contraction is performed. This is the end of the way:

Theorem 5.1. *For each $f \in \mathbf{E}$, there is a limit f_0 to $\mathbf{C}_t f$ for the topology of uniform convergence on compact subsets of $\mathfrak{p}$, and when f belongs to $\mathbf{H}$ this limit belongs to $\mathbf{H}_0$. Moreover, if f_{min} is the orthogonal projection of f onto the lowest K-type isotypical component of $\mathbf{H}$, then $\mathbf{C}_t(f - f_{min})$ tends to zero uniformly on compact sets of $\mathfrak{p}$.*

Remark. The limit f_0 is the constant function on $\mathfrak{p}$ with value $f(0) \in W$.

Let me return to the statement of Theorem 3.2. The space $\mathbf{E} = C(\mathfrak{p}, W)$ of continuous functions from $\mathfrak{p}$ to W is a Fréchet space when equipped with the topology of uniform convergence on compact subsets of $\mathfrak{p}$. What we just saw is that parts 1. to 3. of Theorem 3.2 hold as soon as $\mathbf{M}(\delta)$ is a discrete series representation. To prove parts 4. and 5., we just need the following simple observation.

Lemma 5.6. *Choose $g_0 \in G_0$. Then there is a distance on $\mathbf{E}$ whose associated topology is that of uniform convergence on compact subsets of $\mathfrak{p}$, and with respect to which each of the $\pi_t(\alpha_t g_0)$ is 1-Lipschitz.*

Proof. Whenever $A \subset \mathfrak{p}$ is compact, the subset $\Pi(A) = \{(\alpha_t g_0) \cdot_t A \mid t \in [0,1]\}$ is compact too. So there is an increasing family, say (A_i), of compact subsets of $\mathfrak{p}$, such that $\Pi(A_i) \subset A_{i+1}$, and $\cup_n A_n = \mathfrak{p}$.

A consequence is that for each f and f' in $\mathbf{E}$, $\|\pi_t f - \pi_t f'\|_{A_n} \leq \|f - f'\|_{A_{n+1}}$. Recall that a distance whose associated topology is that of uniform convergence on compact subsets is $d(f,f') = \sum_n \frac{\|f-f'\|_{A_n}}{2^n(1+\|f-f'\|_{A_n})}$. Then $d/2$ has the desired property. □

The proof of Theorem 3.2 for discrete series representations is completed by the next result.

Corollary. *Suppose f is in $\mathbf{E}$, then if $g_0 = (k, v)$, $\pi_t(\alpha_t g_0) f_t$ goes to $\mu(k) f_0$ as t goes to zero.*

Proof. Note first that

$$\pi_t(\alpha_t g_0) f_t = \pi_t(\alpha_t g_0)(f_t - f_0) + \pi_t(\alpha_t g_0) f_0.$$

Because of Lemma 5.6 the first term goes to zero, and because f_0 is a constant function, $\pi_t(\alpha_t g_0) f_0$ is just $\mu(k) f_0$. □

6 Other representations with real infinitesimal character

6.1 Limits of discrete series

If the highest weight $\vec{\mu}$ is integral and C-dominant but the corresponding Harish-Chandra parameter is singular (see (5.1) above), it is no longer true that $\mathbf{V}_G(\mu)$ belongs to the discrete series . But when $\vec{\mu}$ is "not too degenerate", we can build the carrier space for $\mathbf{V}_G(\mu)$ from that of a discrete series representation, following Zuckerman's translation principle: let us consider an element $\vec{\lambda}$ of $\mathfrak{t}^\star$ which is integral, C-dominant and nonsingular. Then we can start from the infinite-dimensional space [9] $\mathbf{H}_{\lambda+\mu}$, which carries a discrete series representation, and form the tensor product $E = \mathbf{H}_{\lambda+\mu} \otimes A(\mu)$, where $A(\mu)$ is the finite-dimensional carrier space of an irreducible representation of $\mathfrak{g}_{\mathbf{C}}$ with lowest weight $-\mu$. Then we can consider the isotypical component

$$E_\lambda := \{v \in E \mid \forall X \in \mathbf{Z}(\mathfrak{g}_{\mathbb{C}}),\ X \cdot v = \xi_{\mathfrak{h}}(\lambda)(X) \cdot v\}.$$

(here $\xi_{\mathfrak{h}}(\lambda)$ is the infinitesimal character of section 3.1).

Zuckerman and Knapp proved ([24], theorem 1.1) that E_λ is an irreducible $(\mathfrak{g}_{\mathbf{C}}, K)$-module and that it has an invariant hermitian form. Depending on μ, this space is either zero or infinite-dimensional; when it is nonzero, it is possible to complete it into an unitary irreducible representation of G, and when we do so the representation is of class $\mathbf{V}_G(\mu)$. If it is not in the discrete series, then it is called a limit of discrete series.

In general it is not easy to describe the unitary structure (think of the explicit, but not easily generalized, Hilbert space norm in the case $SL_2(\mathbf{R})$, see [22], II.5), but after all we shifted the attention away from the Hilbert space norm in these notes; as we shall see the contraction maps $\mathbf{C}_t$ and the weak convergence with respect to the Fréchet topology on E_λ inherited from that of $\mathbf{H}_{\lambda+\mu}$ are not difficult to describe.

We first need to understand how $\mathfrak{g}_C$ acts on the finite-dimensional part $A(\mu)$, and how things evolve when we consider it as a $\mathfrak{g}_{t,\mathbb{C}}$-module. For this, we need to recall a construction for $A(\mu)$. Instead of describing it through its lowest weight, I will write $\tilde{\mu}$ for its highest weight and recall a construction for $A(\mu)$ as the irreducible representation with highest weight $\tilde{\mu}$ (see [23], V.3).

Let us start with the subalgebra

$$\mathfrak{b} = \mathfrak{t}_{\mathbb{C}} \oplus \mathfrak{n} := \mathfrak{t} \oplus \bigoplus_{\alpha \in \Delta^+} \mathfrak{g}_\alpha$$

of $\mathfrak{g}_{\mathbb{C}}$. Setting $\chi(H+E) = (\tilde{\mu} - \rho)(H)$ when H is in $\mathfrak{t}_{\mathbb{C}}$ and E in $\mathfrak{n}$, we obtain an abelian character of $\mathfrak{b}$, and thus an abelian character of the enveloping algebra $U(\mathfrak{b})$. I will write $\mathbb{C}_\chi$ for $\mathbb{C}$ with this $U(\mathfrak{b})$-module structure.

The Verma module $B(\mu)$ is then defined as the induced module

$$B(\mu) = U(\mathfrak{g}_C) \otimes_{U(\mathfrak{b})} \mathbb{C}_\chi.$$

In our setting this means that as a vector space $B(\mu)$ is the quotient $U(\mathfrak{g}_{\mathbb{C}})/M$, with

$$M = \langle Y - \chi(Y), Y \in U(\mathfrak{b}) \rangle$$

(the ideal generated by the $Y - \chi(Y)$s), and that the $U(\mathfrak{g}_{\mathbb{C}})$-action is just the adjoint action passed through the quotient.

Note that $\mathfrak{g}_{\mathbb{C}} = \mathfrak{b} \oplus \mathfrak{n}^-$, with

$$\mathfrak{n}^- := \bigoplus_{\alpha \in \Delta^+} \mathfrak{g}_{-\alpha},$$

and that a consequence is that $U(\mathfrak{n}^-)$, viewed as a vector subspace of $U(\mathfrak{g}_{\mathbb{C}})$, is an algebraic complement to M; the projection from $U(\mathfrak{g}_{\mathbb{C}})$ to $B(\mu)$ restricts to a vector space isomorphism, obviously also a $U(\mathfrak{n}^-)$-module isomorphism, between $U(\mathfrak{n}^-)$ and $B(\mu)$.

Now set

$$S = \text{Sum of all proper submodules of } B(\mu).$$

9. For convenience I will be dropping the "vector" arrows for convenience from now on.

This is of course a submodule, and because the image of 1 (the unit of $U(\mathfrak{g}_{\mathbb{C}})$) in $B(\mu)$ can be contained in no submodule it is actually proper. The irreducible module $A(\mu)$ is the quotient $B(\mu)/S$, and an important step in the classification of finite-dimensional representations is proving that $\dim(A(\mu))$ is finite.

⋆

We know that the isomorphism φ_t extends to an isomorphism $\tilde{\varphi}_t$ between $\mathbf{U}(\mathfrak{g}_{t,\mathbb{C}})$ and $\mathbf{U}(\mathfrak{g}_{\mathbb{C}})$. If $\rho : \mathfrak{g} \to \mathrm{End}(U(\mathfrak{g}_{\mathbb{C}}))$ and $\rho_t : \mathfrak{g}_t \to \mathrm{End}(U(\mathfrak{g}_{t,\mathbb{C}}))$ code for the canonical extensions of the adjoint actions in each of those Lie algebras, then of course $\tilde{\varphi}_t$ intertwines them :

$$\tilde{\varphi}_t \circ \rho_t = \rho \circ \varphi_t.$$

Naturally $\mathfrak{g}$ and $\mathfrak{g}_t$ are the same as vector spaces, so $U(\mathfrak{g}_{t,\mathbb{C}})$ and $U(\mathfrak{g}_{\mathbb{C}})$ are the same as vector spaces too. The construction above applies to $\mathfrak{g}_t$, yielding a Verma module $B_t(\mu) = U(\mathfrak{g}_{t,\mathbb{C}})/(\tilde{\varphi}_t M)$ and a finite-dimensional vector space $A_t(\mu) = B_t(\mu)/S_t$ with natural $\mathfrak{g}_t$-actions intertwined by $\tilde{\varphi}_t$.

Now let S be the sum of all proper $U(\mathfrak{g}_{\mathbb{C}})$ submodules of $B(\mu)$, and ψ_t the map between $U(\mathfrak{g}_{t,\mathbb{C}})/(\varphi_t M)$ and $U(\mathfrak{g}_{\mathbb{C}})/M$ induced by $\tilde{\varphi}_t$. The image $\psi_t S$ is the sum of all proper $U(\mathfrak{g}_{t,\mathbb{C}})$-submodules of $B_t(\mu)$. To study the way vectors in $B(\mu)/S$ evolve as the contraction is performed, we need a way to relate $B_t(\mu)/(\psi_t S)$ with $B(\mu)/S$ inside a fixed space. For this it would be very nice if M and S were invariant, as vector spaces, under the contraction. While I have not been able to see whether it is true that neither M nor S move as the contraction is performed, the next lemma gives a way to view $B_t(\mu)/(\psi_t S)$ as a fixed subspace.

Lemma 6.1.

a. The vector subspace $U(\mathfrak{n}^-)$ of $U(\mathfrak{g}_{\mathbb{C}})$ is an algebraic complement to M which is $\tilde{\varphi}_t$-invariant for all t, so each $\tilde{\varphi}_t$ induces an element, say $\bar{\varphi}_t$, of $GL(B(\mu))$.

b. The maximal proper submodule S of $B(\mu)$ admits an algebraic complement which is $\bar{\varphi}_t$-invariant for all $t > 0$.

Proof.

a. There is an important remark to be made here: because $\mathfrak{t}$ is contained in $\mathfrak{k}$, the real parts of the root spaces for roots of $(\mathfrak{g}_{\mathbb{C}}, \mathfrak{t}_{\mathbb{C}})$ are contained either in $\mathfrak{k}$ (the corresponding roots are called compact roots) or in $\mathfrak{p}$ (the corresponding roots are called noncompact rots). A consequence of this is that as vector subspaces of $\mathfrak{g}$, they will not move during the contraction.

This remark extends to $U(\mathfrak{g}_{\mathbb{C}})$ as follows. There is a natural basis for $\mathbf{U}(\mathfrak{g}_{\mathbb{C}})$ associated to any basis of $\mathfrak{g}$ by the Poincaré-Birkhoff-Witt construction. Let's then choose a basis $(K_1, ...K_n, P_1, ...P_{2q})$ with the K_is in $\mathfrak{k}$ and the P_js in $\mathfrak{p}$, such that a subset of the K_is, say $K_1, ...K_{r_1}$, spans $\mathfrak{t} \oplus \bigoplus_{\alpha \in \Delta_c^+} \mathfrak{g}_\alpha$, a subset of the P_js, say $P_1, ..P_{r_2}$, spans $\bigoplus_{\alpha \in \Delta^+ - \Delta_c^+} \mathfrak{g}_\alpha$, and the other elements span the root spaces for negative roots (so that $r_2 = q$, and n is $2r_1$ plus the rank of $\mathfrak{g}$).

Note first that each element of the associated basis of $U(\mathfrak{g}_{\mathbb{C}})$ is a product $K_1^{u_1}...K_n^{u_n}P_1^{v_1}...P_{2q}^{v_{2q}}$, with $(u_1, ...u_n, v_1, ...v_{2q}) \in \mathbb{N}^{n+2q}$, and that the definition of $\tilde{\varphi}_t$ is equivalent with the fact that

$$\tilde{\varphi}_t \left[K_1^{u_1}...K_n^{u_n} P_1^{v_1}...P_{2q}^{v_{2q}}\right] = t^{v_1+...+v_{2q}} \left[K_1^{u_1}...K_n^{u_n} P_1^{v_1}...P_{2q}^{v_{2q}}\right].$$

The elements of $U(\mathfrak{n}^-)$, viewed as elements of $U(\mathfrak{g}_{\mathbb{C}})$, are just the combinations of those basis elements which have $u_1 = ... = u_{r_1} = v_1 = ... = v_{r_2} = 0$. So the subspace $U(\mathfrak{n}^-)$ of $U(\mathfrak{g}_C)$ is indeed $\tilde{\varphi}_t$-invariant for all t.

b. The second part is a consequence of the following simple observation:

Lemma 6.2. *Suppose $V = \bigoplus_{k\geq 0} V^k$ is a graded vector space, and S is a linear subspace with finite codimension. Then there is an algebraic complement to S for which a basis consists of homogeneous elements.*

Proof. I will write $v^{\max}$ for the highest-degree homogeneous component of a vector v in V here.

Let's use induction on the codimension of S. If $\mathrm{codim}(S)$ is one, and $V = \mathbb{C}e_1 \oplus S$, it is not possible that every homogeneous component of e_1 be in S. Any homogeneous component that is not in S then yields a homogeneous algebraic complement to S.

Suppose now $\mathrm{codim}(S)$ is higher. When E is a finite-dimensional subspace of V, let me write $d_E = \max\left(d \in \mathbb{N} \,|V^d \cap E \neq \{0\}\right)$. Choose d as the smallest integer such that there is an algebraic complement E to S with $d_E = d$, and let me start with E_0 such that $E_0 \oplus S = V$ and $d_{E_0} = d$. Choose a basis $(e_1, ...e_n)$ of E_0, and order it so that the e_is have decreasing degrees, and $(e_1, ...e_k)$ are the ones with maximal degree. Then there are two possible cases :

Case 1: $e_1^{max} \notin Span\left[e_2, ...e_n\right] \oplus S$. Then $V = \mathrm{Span}\left[e_2, ...e_n\right] \oplus (\mathbb{C}e_1 \oplus S)$, and the conclusion for S follows from the induction hypothesis.

Case 2: $e_1^{max} \in Span\left[e_2, ...e_n\right] \oplus S$. Then $E_1 = \mathrm{Span}\left[e_2, ...e_k, e_1 - e_1^{\max}, e_{k+1}, ...e_n\right]$ is an algebraic complement to S. Check whether $e_2^{\max}$ is in $\mathrm{Span}\left[e_3, ...e_k, e_1 - e_1^{\max}, e_{k+1}, ...e_n\right] \oplus S$, and if it is, define $E_2 = \mathrm{Span}\left[e_2, ...e_k, e_2 - e_2^{\max}, e_1 - e_1^{\max}, e_{k+1}, ...e_n\right]$ and start again. This algorithm cannot fail to produce a situation in which Case 1 appears for one E_i, $i \leq k$, since if that were the case d would not be minimal. Lemma 6.2 follows. ☐

To prove lemma 6.1.b, we use the grading on $B(\mu)$ provided by the isomorphism between $U(\mathfrak{n}^-)$ and $B(\mu)$, deciding that the image of $\left[K_{r_1+1}^{u_{r_1+1}}...K_n^{u_n} P_{r_2+1}^{v_{r_2+1}}...P_{2q}^{v_{2q}}\right]$ in $B(\mu)$ has degree $v_{q+1} + ... + v_{2q}$. The linear map $\bar{\varphi}_t$ then acts as multiplication by t^v on the subspace consising of homogeneous elements with degree v, so that a subspace generated by homogeneous elements is $\bar{\varphi}_t$-stable. We can then use lemma 6.2 to conclude the proof of lemma 6.1.

☐

Because of lemma 6.1, we know that there is a $\tilde{\varphi}_t$-invariant, finite-dimensional subspace $F(\mu)$ of $U(\mathfrak{n}^-)$ on which for each t, the composition of the two projections from $U(\mathfrak{g}_{t,\mathbb{C}})$ to $B_t(\mu)$ and from $B_t(\mu)$ to $A_t(\mu)$ restricts to a linear isomorphism. We know that φ_t induces a linear map which intertwines the actions of $\mathfrak{g}$ and $\mathfrak{g}_t$ on $A(\mu)$ and $A_t(\mu)$, so using our linear isomorphisms to lift these actions to $F(\mu)$, we end up with maps ρ' and ρ'_t from $\mathfrak{g}$ and $\mathfrak{g}_t$ to $\mathrm{End}(F_\mu)$, which turn F_μ into a finite-dimensional irreducible $\mathfrak{g}_{\mathbb{C}}$-module with lowest weight $-\mu$ and a finite-dimensional irreducible $\mathfrak{g}_{t,\mathbb{C}}$-module with lowest weight $-\mu$, and which satisfy in addition

$$\tilde{\varphi}_t \circ \rho'_t = \rho' \circ \varphi_t.$$

We have thus exhibited our linear map $\tilde{\varphi}_t$ as a contraction map from $F(\mu)$ to itself. We now rename it as $\mathbf{C}_t^{fd}$.

But we explicitly know how $\tilde{\varphi}_t$ acts on $\mathbf{U}(\mathfrak{g}_C)$, so we can use this to see whether there is a limit to this contraction operator as t goes to zero. In the proof of lemma 6.1.b, we saw that a linear basis for $U(\mathfrak{g}_C)$ consists of monomials for which

$$\tilde{\varphi}_t \left[K_1^{u_1}...K_n^{u_n}P_1^{v_1}...P_{2q}^{v_{2q}}\right] = t^{v_1+...+v_{2q}} \left[K_1^{u_1}...K_n^{u_n}P_1^{v_1}...P_{2q}^{v_{2q}}\right].$$

But of course these formulae make sense in the limit $t = 0$. Here is the conclusion :

Lemma 6.3. *For each $v \in F(\mu)$, there is a limit to $\mathbf{C}_t^{fd}v$ as t goes to zero.*

Here the convergence is in the sense of any norm-induced topology on $A(\mu)$, and the limit is naturally an element of $U(\mathfrak{k}_\mathbb{C})$.

It is time to return to limits of discrete series. Suppose $\mathbf{E}^{ds}_{\lambda+\mu}$ is the Fréchet space we associated to the representation $\mathbf{H}_{\lambda+\mu}$ in section 5. Consider now a vector F in $\mathbf{E} = \mathbf{E}^{ds}_{\lambda+\mu} \otimes F(\mu)$. It can be written as a finite sum $F = \sum_i f_i \otimes v_i$, with the f_is in $\mathbf{E}^{ds}_{\lambda+\mu}$ and the v_is in $A(\mu)$. We now set

$$\mathbf{Z}_t f = f_t := \sum_i \left(\mathbf{C}_t^{ds} f_i\right) \otimes \left(\mathbf{C}_t^{fd} v_i\right)$$

where $\mathbf{C}_t^{ds} \in \mathrm{End}\left(\mathbf{E}^{ds}_{\lambda+\mu}\right)$ is the contraction operator defined in section 5.2.

Lemma 6.3, together with the results of section 5, yields :

Lemma 6.4. *For each vector $F \in \mathbf{E}$, there is a limit F_0 to $\mathbf{Z}_t F$ as t goes to zero.*

Let us now see what remains if we start from the carrier Hilbert space $\mathbf{E}_\lambda$ of our limit of discrete series, viewed as a vector subspace of $\mathbf{E}$.

For the moment the map $\mathbf{Z}_t$ is defined on all of $\mathbf{E}$, which is much larger than the space we are actually interested in. If this $\mathbf{Z}_t$ is to be our contraction map between representation spaces, we need the following fact. Let me use the notations of section 5.2 and write $\mathbf{E}_\lambda$ for the vector subspace $\mathbf{H}_{\lambda+\mu} \otimes F(\mu)$ of $\mathbf{E}$, which carries our limit of discrete series representation as recalled above, and $\mathbf{E}_{t,\lambda}$ for the vector subspace $\mathbf{H}^t_{\lambda+\mu} \otimes F(\mu)$.

Lemma 6.5. *For each $F \in \mathbf{E}_\lambda$, $\mathbf{Z}_t F$ belongs to $\mathbf{E}_{t,\lambda}$*

To prove this, we need to start with an element $X \in \mathbf{Z}(\mathfrak{g}_{t,\mathbb{C}})$ and to see how it acts on $\mathbf{Z}_t F$. What we know is the infinitesimal character of the action of $U(\mathfrak{g}_\mathbb{C})$ on $\mathbf{E}_\lambda$, so writing π and π_t for the actions of $\mathfrak{g}_\mathbb{C}$ and $\mathfrak{g}_{t,\mathbb{C}}$ on $\mathbf{E}$ naturally defined from those in Section 5, we know that $\pi(\tilde{\varphi}_t X)F$ is $\xi_\mathfrak{t}(\lambda)(\tilde{\varphi}_t X)F$. Because $\mathbf{Z}_t$ intertwines the actions on $\mathbf{E}$ by definition, this means that $\pi_t(X)(\mathbf{Z}_t F) = \xi_\mathfrak{t}(\lambda)(\tilde{\varphi}_t X)F$.

Does this mean that $X \mapsto \xi_\mathfrak{t}(\lambda)(\tilde{\varphi}_t X)$ is the abelian character of $\mathbf{Z}(\mathfrak{g}_{t,\mathbb{C}})$ which, through the Harish-Chandra isomorphism associated to the pair $(\mathfrak{g}_{t,\mathbb{C}}, \mathfrak{t}_\mathbb{C})$, has parameter λ ? Yes, it does.

For this I recall the definition of the Harish-Chandra isomorphism γ between $\mathbf{Z}(\mathfrak{g}_{\mathbb{C}})$ and $\mathfrak{t}^\star/W(\mathfrak{g}_{\mathbb{C}}, \mathfrak{t}_{\mathbb{C}})$ (see [23], V.7): one starts with the decomposition $\mathfrak{g}_{\mathbb{C}} = \mathfrak{t}_{\mathbb{C}} \oplus \mathfrak{n} \oplus \mathfrak{n}^-$, and this yields a direct sum decomposition

$$U(\mathfrak{g}_{\mathbb{C}}) = U(\mathfrak{t}_{\mathbb{C}}) \oplus \left[U(\mathfrak{g}_{\mathbb{C}})\mathfrak{n} \oplus \mathfrak{n}^- U(\mathfrak{g}_{\mathbb{C}})\right].$$

Write $p_{U(\mathfrak{t}_{\mathbb{C}})}$ for the associated projection $U(\mathfrak{g}_{\mathbb{C}}) \to U(\mathfrak{t}_{\mathbb{C}})$, and recall that ρ is the half-sum of positive roots of $(\mathfrak{g}_{\mathbb{C}}, \mathfrak{t}_{\mathbb{C}})$ with respect to the ordering we have been working with in this section. The linear map

$$H \in \mathfrak{t}_{\mathbb{C}} \mapsto H - \rho(H)1 \in U(\mathfrak{t}_{\mathbb{C}})$$

extends to an algebra automorphism, say τ, of $U(\mathfrak{t}_{\mathbb{C}})$, and the Harish-Chandra isomorphism is

$$\gamma := \tau \circ p_{U(\mathfrak{t}_{\mathbb{C}})}.$$

Of course this construction also yields an algebra isomorphism γ_t between $\mathbf{Z}(\mathfrak{g}_{t,\mathbb{C}})$ and $\mathfrak{t}^\star/W(\mathfrak{g}_{\mathbb{C}}, \mathfrak{t}_{\mathbb{C}})$, and I claim that $\gamma_t = \gamma \circ \tilde{\varphi}_t$. The reason is that

If α is a root of $(\mathfrak{g}_{\mathbb{C}}, \mathfrak{t}_{\mathbb{C}})$, then it is also a root of $(\mathfrak{g}_{t,\mathbb{C}}, \mathfrak{t}_{\mathbb{C}})$, and the root spaces correspond under ϕ_t.

Indeed, if X is an element of $\mathfrak{g}_{\mathbb{C}}$ such that $[H, X]_{\mathfrak{g}_{\mathbb{C}}} = \alpha(H)X$ for all H in $\mathfrak{t}_{\mathbb{C}}$, then $[\phi_t^{-1}H, \phi_t^{-1}X]_{\mathfrak{g}_{t,\mathbb{C}}} = \alpha(H)\phi_t^{-1}X$, and the statement in italics follows because ϕ induces the identity on $\mathfrak{t}_{\mathbb{C}}$ (this is the difference with Lemma 4.2).

An immediate consequence is that γ_t is defined from the decomposition $\mathfrak{g}_{t,\mathbb{C}} = \mathfrak{t}_{\mathbb{C}} \oplus (\phi_t^{-1}\mathfrak{n}) \oplus (\phi_t^{-1}\mathfrak{n}^-)$, hence that the projection $p_{t,U(\mathfrak{t}_{\mathbb{C}})}$ defined from $\mathfrak{g}_t$ is just $p_{U(\mathfrak{t}_{\mathbb{C}})} \circ \tilde{\varphi}_t$. Another immediate consequence is that the half-sum of positive roots of $(\mathfrak{g}_{t,\mathbb{C}}, \mathfrak{t}_{\mathbb{C}})$ is also ρ, so that the shift between $p_{t,U(\mathfrak{t}_{\mathbb{C}})}$ and γ_t is still τ. This proves lemma 6.5. □

Because I defined $\mathbf{Z}_t$ in a manner compatible with the definition of tensor products of $\mathfrak{g}_{\mathbb{C}}$-modules, and because of Lemma 6.5, we now know that $\mathbf{Z}_t$ intertwines the actions of $\mathfrak{g}_{\mathbb{C}}$ and $\mathfrak{g}_{t,\mathbb{C}}$ on $\mathbf{E}_\lambda$ and $\mathbf{E}_{\lambda,t}$. When these infinitesimal versions are integrated and $\mathbf{E}_{\lambda,t}$ is viewed as the space of smooth, K-finite vectors in the carrier space of a unitary irreducible representation of G_t, $\mathbf{Z}_t$ becomes a well-defined contraction operator in the sense of section 3.3; we now rename $\mathbf{Z}_t$ as $\mathbf{C}_t$.

Lemma 6.6. *The vector space $\mathbf{H}_0 := \{F_0 \mid F \in E_\lambda\}$ carries an irreducible K-module of class μ.*

Proof. Let me write $\mathbf{H}^0_{\lambda+\mu}$ for the finite-dimensional vector space gathering the limits of the $\mathbf{C}_t^{ds} f$, $f \in \mathbf{H}_{\lambda+\mu}$, and $F(\mu)^0$ for the subspace of $F(\mu)$ gathering the limits of the $\mathbf{C}_t^{fd} v$, $v \in F(\mu)$. Our $\mathbf{H}_0$ is then the image under p_λ of the tensor product $\mathbf{H}^0_{\lambda+\mu} \otimes F(\mu)^0$, viewed as a subspace of $\mathbf{E}$.

But suppose we start with the tensor product, say $\mathbf{A}^0_{\lambda+\mu} \otimes G(\mu)^0$, between the carrier space for an irreducible K-module with highest weight $\lambda + \mu$ and the carrier space for an irreducible $\mathfrak{k}_{\mathbb{C}}$-module with lowest weight $-\mu$, then look at the isotypical component corresponding to the infinitesimal character which the Harish-Chandra isomorphism for the pair $(\mathfrak{k}_{\mathbb{C}}, \mathfrak{t}_{\mathbb{C}})$ associates to λ. Then, because K is a reductive Lie group and an irreducible K-module with highest weight $\lambda + \mu$ can be viewed as a discrete series representation, a trivial case of the result described above for limits of discrete series says that this isotypical component is the carrier space for an irreducible representation with highest weight λ.

Now, it is true that $F(\mu)^0$ is the carrier space for an irreducible $\mathfrak{k}_{\mathbb{C}}$-module with lowest weight $-\mu$: the definition of $F(\mu)$ means that

$$F(\mu)^0 = F(\mu) \cap U(\mathfrak{k}_{\mathbb{C}}) \cap U(\mathfrak{n}^-) = F(\mu) \cap \sum_{\alpha \in \Delta_c^+} \mathfrak{k}_{-\alpha};$$

and except at zero $F(\mu)^0$ does not intersect M, especially not

$$M^0 := M \cap U(\mathfrak{k}_{\mathbb{C}}) = \left\langle Y - \chi(Y), Y \in U(\mathfrak{t}_{\mathbb{C}} \oplus \sum_{\alpha \in \Delta_c^+} \mathfrak{k}_\alpha) \right\rangle.$$

In addition, the image of $F(\mu)^0$ in $U(\mathfrak{k}_{\mathbb{C}})/M^0$ is an algebraic complement to

$$S^0 := S \cap U(\mathfrak{k}_{\mathbb{C}}) = \text{Sum of all proper } U(\mathfrak{k}_{\mathbb{C}}) \text{ submodules of } U(\mathfrak{k}_{\mathbb{C}})/M^0$$

(the last equality is because the proper $U(\mathfrak{g}_{\mathbb{C}})$-submodules of $B(\mu)$ are those that do not contain the highest weight $\tilde{\mu} - \rho$ upon restriction to $\mathfrak{t}$, and each of those decomposes under $U(\mathfrak{k}_{\mathbb{C}})$ as a sum of $U(\mathfrak{k}_{\mathbb{C}})$-submodules which do not contain the highest weight $\tilde{\mu} - \rho_c$ upon restriction to $\mathfrak{t}$, so that they project in $U(\mathfrak{k}_{\mathbb{C}})/M^0$ as proper $U(\mathfrak{k}_{\mathbb{C}})$-modules).

Hence the double projection map from $F(\mu)^0$ to $(U(\mathfrak{k}_{\mathbb{C}})/M^0)/S^0$ is a vector space isomorphism which commutes with the action of $\mathfrak{k}$ on both spaces, as announced.

This proves lemma 6.6. □

Let me summarize the situation for limits of discrete series representations: because of lemma 6.6 and the interpretation of the "limit subspace" for discrete series as a space of constant functions, it is also true that our Fréchet space $\mathbf{E}_\lambda$ can be viewed as the space of continuous functions with values in a fixed vector space carrying the minimal K-type of our representation π. The convergence of vectors in the subspace of $\mathbf{E}_\lambda$ which carries the limit of discrete series is summarized in the following statement.

Theorem 6.1. *For each $f \in \mathbf{E}$, there is a limit f_0 to $\mathbf{C}_t f$ for the Fréchet topology of uniform convergence on compact subsets of $\mathfrak{p}$ when one views $\mathbf{E}$ as a space of $W_{\lambda+\mu} \otimes F(\mu)$-valued functions on $\mathfrak{p}$; the space of limits obtained from elements of $\mathbf{E}_\lambda$ carries an irreducible K-module whose equivalence class is the lowest K-type of π. Moreover, if f_{min} is the orthogonal projection of f onto the lowest K-type isotypical component of $\mathbf{H}$, then $\mathbf{C}_t(f - f_{min})$ tends to zero uniformly on compact sets of $\mathfrak{p}$.*

Note that because $F(\mu)$ is a priori larger than $F(\mu)^0$, when we interpret $\mathbf{E}_\lambda$ as a space of functions on $\mathfrak{p}$ the value of functions at zero is modified by the contraction process: in the limit it is projected on $W \otimes F(\mu)^0$. This did not happen in the case of discrete series. To come back to the statement of Theorem 3.2, it is no longer quite true that $\mathbf{C}_t$ is the only contraction map which preserves the value of functions at zero, but it is the only contraction map which preserves the projection on $W \otimes F(\mu)^0$ of their value at zero.

Because the action of $\mathfrak{g}$ on the finite-dimensional part is through bounded operators, the end of point 4. in Theorem 3.2 follows immediately from the analogous statement for the discrete series (lemma 5.6 and the corollary), and this completes the proof of Theorem 3.2 when $\mathbf{M}(\delta)$ is a nonzero limit of discrete series.

6.2 Real-infinitesimal-character representations which are not limits of discrete series

We will now consider the representations of G which are irreducible tempered and have real infinitesimal character, and hence a minimal K-type, but which are neither in the discrete series nor limits of discrete series.

Example 6.1. *When G is $SL_2(\mathbf{R})$, there is only one such representation: the irreducible principal series representation with continous parameter zero, whose minimal K-type is the trivial representation.*

For the representation $\mathbf{V}_G(1)$ whose minimal K-type is the trivial representation, we gave two geometric realizations in section 4: in the compact picture, the Hilbert space is $\mathbf{L}^2(K)$ and $g \in G$ acts as $f \mapsto \left[k \mapsto \exp\langle -\rho, \mathbf{a}(g^{-1}k)\rangle \sigma(\mathbf{m}(g^{-1}k)) f\left(\kappa(g^{-1}k)\right)\right]$, and in the second the Hilbert space is a space of functions on $\mathfrak{p}$.

6.2.1. The trivial representation of G_0 in Helgason's picture.
Let us start with Helgason's picture, and recall that in section 4.2. I used functions $e_{\lambda,b}$ on G/K. The definition makes sense with $\lambda = 0$, and it is true that

$$\mathbf{H}^{Helgason} = \left\{ \int_{K/M} e_{0,b} F(b) db \mid F \in \mathbf{L}^2(K/M) \right\}$$

carries an irreducible representation of G whose equivalence class is $\mathbf{V}_G(1)$. Because of the results in section 4, which do hold when $\lambda = 0$, we know that the contracted waves $\varepsilon^t_{0,b}$ give rise to the corresponding representation of G_t with Hilbert space

$$\mathbf{H}^{Helgason}_t = \left\{ \int_{K/M} \varepsilon^t_{0,b} F(b) db \mid F \in \mathbf{L}^2(K/M) \right\}$$

But now when t goes to zero, all of the $\varepsilon^t_{0,b}$ converge to the constant function with value 1 ! Because of the results of Section 4.2.b, the map $\mathbf{C}_t : f \mapsto [x \mapsto f(tx)]$ turns out to be a contraction map between the various $\mathbf{H}^{Helgason}_t$s (here the renormalization of frequencies has no effect, and does not break the equivariance); the conclusion is that if $\mathbf{E}$ is the space of continuous , complex-valued functions on $\mathfrak{p}$ equipped with the topology of uniform convergence on compact subsets,

Theorem 6.2. *For each $f \in \mathbf{E}$, there is a limit to $\mathbf{C}_t f$ as t goes to zero; the limit is the constant function with value $f(0)$. For each g_0 in G_0 and each f in $\mathbf{H}^{Helgason}$, $\left[x \mapsto (\mathbf{C}_t f)\left((\alpha_t g_0)^{-1} \cdot_t x\right)\right]$ converges (in $\mathbf{E}$) to the constant function with value $f(0)$.*

This proves Theorem 3.2 when $\mathbf{M}_0(\delta)$ is the trivial representation.

6.2.2. The trivial representation of G_0 in the compact picture.
If instead of Helgason's picture we take up the compact picture and try to perform the contraction, the situation is less promising. Here $\mathbf{H}^{comp}$ is just $\mathbf{L}^2(K/M)$, and as we saw earlier, the only contraction map between $\pi^{comp}_{0,1}$ and $\pi^{t,comp}_{0,1}$ which preserves the value of functions at zero is the identity ! So in the limit, we will certainly not get the carrier space for the trivial representation. Instead, the proof of theorem 4.1 shows that if $g_0 = (k, v)$, $\pi^{t,comp}_{\lambda,\sigma}(\alpha_t(k,v))$ weakly converges to $f \mapsto \left[u \mapsto f(k^{-1}u)\right]$, so that in the limit we get the quasi-regular representation of K on $\mathbf{L}^2(K/M)$ instead of the trivial representation of K !

6.2.3. A remark on the other representations.

It would be very nice if the other real-infinitesimal-character cases could be treated in the same way. I do not see how, though. The only description I know for tempered irreducible representations which have real infinitesimal character, but are not in the discrete series or limits of discrete series, is a simple consequence of the Knapp-Zuckerman classification theorem, already alluded to at the end of section 3.2 .

The next statement holds as soon as G satisfies the axioms in section 1 of [24]; linear connected reductive groups do satisfy the axioms, and when $P = M_P A_P N_P$ is a parabolic subgroup in such a group, M_P does, too.

Fact (Knapp-Zuckerman). If σ is a irreducible tempered representation of G which has real infinitesimal character, then there is a *cuspidal* parabolic subgroup MAN of G, and there is a discrete series or nondegenerate limit of discrete series representation $\sigma^\flat$ of M, such that

$$\sigma = Ind_{MAN}^{G}\left(\sigma^\flat \otimes \mathbf{1}\right).$$

I will not need to say precisely what it means to be a "nondegenerate limit of discrete series" (see [24], section 1 and 8, for details). Let me remark here that when M is a reductive group in Harish-Chandra's class *with a nonempty discrete series*, the discrete-series-or-non-degenerate-limit-of-discrete-series representations account for a "large" (but proper) part of the tempered real-infinitesimal-character representations of M.

Given what I said in sections 6.2.1 and 6.2.2, it is pretty clear that the compact picture for σ that the above fact provides will not be enough to get an irreducible K-module as the outcome of the contraction process. Let me be a bit more precise here, although I will need to anticipate on a few results to come, especially section 7.2 below. There I will describe how the compact picture for $Ind_{MAN}^{G}\left(\sigma^\flat \otimes \mathbf{1}\right)$ makes it possible to describe the contraction process in terms quite analogous to Theorem 3.2. However, section 6.2.2 makes it natural to expect (and section 7.2 will prove) that the outcome of the contraction is a *reducible* K-module isomorphic with $Ind_{K\cap M}^{K}(\mu^\flat)$, where $\mu^\flat$ is the minimal $K\cap M$-type of $\sigma^\flat$. This is a bit disappointing of course; I should however mention that $Ind_{K\cap M}^{K}(\mu^\flat)$ contains μ with multiplicity one (and in fact μ is its only minimal K-type, as should be clear from section 2 in Chapter 8 below).

The other usual pictures do not seem to lead to a setting in which the contraction can easily be described. In view of what precedes, the following question seems natural: is there a realization for this which would be analogous to Helgason's picture, and would allow for the Hilbert space for σ to be viewed as a space of functions on $\mathfrak{p}$, or perhaps on a vector subspace or quotient or $\mathfrak{p}$? To my knowledge none has been set forth yet. It is likely that Camporesi's paper [5] might be helpful in that direction, but I have not looked deep enough into the matter at present.

7 General tempered representations

7.1 Discrete series for disconnected groups

To cover the general case, we need to describe the discrete series representations of the (a priori disconnected) M_χ when χ is an element of $\mathfrak{a}^\star$. I shall follow [24] here and refer to [22], XII.8.

Throughout this paragraph, M will be a reductive group satisfying the axioms in section 1 of [24] (as we saw, this is more restrictive than M being in Harish-Chandra's class).

Let us first consider the identity component M_0 of M. It is a non-semisimple, connected Lie group and can be decomposed as $M_0 = M_{ss}\,(Z_M)_0$, with M_{ss} a connected semisimple Lie group with finite center. The abelian group $(Z_M)_0$ is compact and central in M_0 ([22], section V.5).

Suppose we start with a discrete series representation of M_0. Then the elements in $(Z_M)_0$ will act as scalars, and we will get an abelian character of $(Z_M)_0$. It is easy to check that the restriction to M_{ss} of our representation will then be irreducible and belong to the discrete series of M_{ss}, because its matrix elements will be square-integrable.

A discrete series representation π_0 of M_0 is thus specified by a discrete series representation π_{ss} of M_{ss} and an abelian character ξ of $(Z_M)_0$ whose restriction to $M_{ss} \cap (Z_M)_0$ coincides with $(\pi_{ss})\big|_{M_{ss}\cap(Z_M)_0}$. The Hilbert space for π_0 is that of π_{ss}, and the formula for π_0 is $g = g_{ss}g_{(Z_M)_0} \mapsto \xi(g_{(Z_M)_0})\pi_{ss}(g_{ss})$.

Now that we know how to describe the discrete series of M_0, let us write $M^\sharp$ for the subgroup $M_0 Z_M$ of G; because of [22], lemma 12.30, M_0 has finite index in $M^\sharp$, and in addition there is a finite, abelian subgroup F of K (it is the subgroup called $F(B^-)$ in [22]) such that

$$M^\sharp = M_0 F$$

and F is in the center of M (hence of $M^\sharp$).

Of course the arguments we recalled for M_0 go through here, and a discrete series representation $\pi^\sharp$ of $M^\sharp$ is thus specified by a discrete series representation π_0 of M_0 and an abelian character χ of F whose restriction to $M_0 \cap F$ coincides with $(\pi_0)\big|_{M_0\cap F}$. The Hilbert space for $\pi^\sharp$ is that of π_0, and the formula for $\pi^\sharp$ is $g = g_0 f \mapsto \chi(f)\pi_0(g_0)$.

To obtain a unitary representation of M, we can start from a discrete series representation $\pi^\sharp$ of $M^\sharp$ and set

$$\pi = \mathrm{Ind}_{G^\sharp}^{G}\left(\pi^\sharp\right).$$

It turns out ([22], Proposition 12.32) that π is irreducible, is in the discrete series of M, and that $\pi^\sharp \mapsto \pi$ maps the discrete series of $M^\sharp$ onto the discrete series of M. In addition, the restriction of π to $M^\sharp$ decomposes as

$$\pi\big|_{M^\sharp} = \sum_{w\in M/M^\sharp} w\pi^\sharp$$

where $w\pi^\sharp$ is $m \mapsto \pi^\sharp(w^{-1}mw)$. Notice that $M^\sharp$ has finite index in M (see (12.74) in [22]), so the sum is finite here.

⋆

Now, the above description makes it easy to describe the contraction maps between the carrier spaces for discrete series representations of M and M_t.

A first remark is that both $(Z_M)_0$ and F are contained in K (see [22], sections V.5 and XII.8). So we can consider the discrete series representations of M and M_t assembled from χ, ξ and discrete series representations of M_{ss} and $(M_t)_{ss}$ with the same minimal $(K \cap M_{ss})$-type, and these will have the same minimal $(K \cap M)$-type.

Suppose now $\mathbf{H}$ and $\mathbf{H}_t$ are the carrier spaces for discrete series representations of M_{ss} and $(M_{ss})_t$ with the same minimal K_{ss}-type, and suppose $\mathbf{C}_t$ is the contraction map between $\mathbf{H}$ and $\mathbf{H}_t$ defined in Section 5. Then for each choice of ξ and χ, $\mathbf{C}_t$ intertwines the discrete series representations representations of $M^\sharp$ and $(M_t)^\sharp$. This is because the elements of M_0 and F act through multiples of the identity on $\mathbf{H}$, so of course

$$\pi^\sharp \left(g_{ss} g_{(Z_M)_0} f\right) \circ C_t = \xi(g_{(Z_M)_0})\chi(f)\left(\pi^{ss} \circ \mathbf{C}_t\right).$$

Let us then start with discrete series representations of M and M_t which have the same minimal $K \cap M$-type. Let's write the decomposition of their restrictions to $M^\sharp$ and $(M_t)^\sharp$ as

$$\mathbf{H} = \sum_{w \in M/M^\sharp} \mathbf{H}_w$$

$$\mathbf{H}_t = \sum_{w \in M/M^\sharp} \mathbf{H}_{t,w}$$

and suppose we intertwine each of the summands with a geometric realization as a space of solutions of a Dirac equation, so that for each ω, we can view $\mathbf{H}_\omega$ and the various $\mathbf{H}_{t,\omega}$s ($t > 0$) as subspaces of a fixed Fréchet space $\mathbf{E}_\omega$ as described in section 4. We can then view $\mathbf{H}$ and the $\mathbf{H}_t$s, $t > 0$, as subspaces of a fixed Fréchet space $\mathbf{E}$ (the finite direct sum of the $\mathbf{E}_\omega$s). In section 4 we defined maps $\mathbf{C}_t^\omega$ from $\mathbf{E}_\omega$ to itself which send $\mathbf{H}_\omega$ to $\mathbf{H}_{t,\omega}$.

Let us define

$$\mathbf{C}_t \left(\sum_{\omega \in M/M^\sharp} f_\omega \right) = \sum_{\omega \in M/M^\sharp} \mathbf{C}_t^\omega f_\omega.$$

This is a linear map from $\mathbf{E}$ to itself sending $\mathbf{H}$ to $\mathbf{H}_t$.

Lemma 7.1. *This operator is a contraction map.*

Proof. Of course this map is defined in such a way that it commutes with the restrictions to $M^\sharp$ and $(M_t)^\sharp$. What we need to check is just that is commutes with the M and M_t-actions. But this is clear from the definition of induced representations when one induces from a subgroup with finite index: suppose m is in M and $(m_\omega)_{\omega \in M/M^\sharp}$ is a section of the projection $M \to M/M^\sharp$ (so each m_ω is in M), then there is a collection $(m_\omega^\sharp)_{\omega \in M/M^\sharp}$ of elements of $M^\sharp$ such that $m m_\omega = m_{[m\omega]} m^\sharp_{[m\omega]}$, and the action of m on $\mathbf{H}$ is

$$\sum_\omega x_\omega \mapsto \sum_\omega \pi^\sharp_\omega (m^\sharp_{[m\omega]}) x_{[m\omega]}.$$

Then $M_t/M_t^\sharp$ and $M/M^\sharp$ coincide, $(\varphi_t^{-1} m_\omega)_{\omega \in M/M^\sharp}$ is a full set of representatives, $\varphi_t^{-1} m$ will satisfy $(\varphi_t^{-1} m)(\varphi_t^{-1} m_\omega) = (\varphi_t^{-1} m_{[m\omega]})(\varphi_t^{-1} m^\sharp_{[m\omega]})$, and will act on $\mathbf{H}_t$ through

$$\sum_\omega x_\omega \mapsto \sum_\omega \pi^\sharp_{t,\omega} (\varphi_t^{-1} m^\sharp_{[m\omega]}) x_{[m_\omega]}.$$

Then of course

$$\begin{aligned}
\mathbf{C}_t \left(\pi(m) \sum_\omega f_\omega \right) &= \sum_\omega \mathbf{C}_t^\omega \pi^\sharp_\omega (m^\sharp_{[m\omega]}) f_{[m_\omega]}. \\
&= \sum_\omega \pi^\sharp_{t,\omega} (\varphi_t^{-1} m^\sharp_{[m\omega]}) \mathbf{C}_t^\omega f_{[m_\omega]}. \\
&= \pi_t(\varphi_t^{-1} m) \left(\sum_\omega \mathbf{C}_t^\omega f_{[m_\omega]} \right) \\
&= \pi_t(\varphi_t^{-1} m) \left(\mathbf{C}_t \left[\sum_\omega f_\omega \right] \right)
\end{aligned}$$

and the lemma follows. □

Lemma 7.2. *For each vector $F \in \mathbf{E}$, there is a limit F_0 to $\mathbf{C}_t F$ as t goes to zero.*

The limit in the statement is with respect to the Fréchet topology of $\mathbf{E}$, and the lemma is obvious from Theorem 5.1. Now our aim was to describe the contraction of a discrete series representation onto a space carrying a K-module whose equivalence class is the minimal K-type μ of the discrete series we started from, so the following result is the end of the way :

Lemma 7.3. *The vector space $\mathbf{H}_0 := \{F_0 \mid F \in \mathbf{E}\}$ carries an irreducible K-module of class μ.*

Proof. Let me write $K^\sharp$, K_0, K_{ss} for the intersections of K with $M^\sharp$, M_0, M_{ss}. For each $\omega \in M/M^\sharp$, write

$$\mathbf{V}_\omega := \{F_0 \mid F \in \mathbf{H}_\omega\}.$$

This is an irreducible K_{ss}-module whose equivalence class is the minimal K_{ss}-type, say $\mu_\omega^\flat$, of $\mathbf{H}_\omega$. One can use the characters ξ and χ to turn $\mathbf{V}_\omega$ into a $K^\sharp$-module as above; I will write $\mu_\omega^\sharp$ for its equivalence class, which is also the minimal $K^\sharp$-type of $\mathbf{H}_\omega$.

Now, we know that the inclusion from K to M induces an isomorphism between $K/K^\sharp$ and $M/M^\sharp$ (see (12.74) in [22]), so the outcome of the contraction can be rewritten as

$$\mathbf{H}_0 = \sum_{K/K^\sharp} \mathbf{V}_\omega.$$

The fact that each $\mathbf{C}_t^\omega$ is K-equivariant and induces an intertwining map between the restriction of π^ω to the minimal K-type component of $\mathbf{H}_\omega$ on the one hand, and the action μ^ω on $\mathbf{V}_\omega$ on the other hand, means that the action of K on $\mathbf{H}$ will induce an action of K on $\mathbf{H}_0$. In this way an element k in K will act as

$$\sum_\omega x_\omega \mapsto \sum_\omega \mu_\omega^\sharp(k^\sharp_{[k\omega]}) x_{[k\omega]}$$

where the $k^\sharp_{[k\omega]}$s are the elements defined in the proof of lemma 7.1 if we take care to ask that the representatives m_ω, $\omega \in M/M^\sharp$, belong to K.

Of course the description of induced representations given in the proof of lemma 7.1 means that for any ω_0 in $K/K^\sharp$,

$$\mathbf{H}_0 \simeq \mathrm{Ind}_{K^\sharp}^K \left(\mu_{\omega_0}^\sharp\right).$$

But as a particular case of the description of the discrete series of a reductive group M from the discrete series of M^{ss} recalled above, we do know that

$$\mu \simeq \mathrm{Ind}_{K^\sharp}^K \left(\mu_{\omega_0}^\sharp\right).$$

So the equivalence class of $\mathbf{H}_0$ as a K-module is really that of μ. □

I worked with discrete series representations of M_χ for convenience here, but it is clear from the constructions recalled above that the remarks in this subsection yield a description of both the limits of discrete series representations of M_χ and their contraction onto their minimal K-type.

7.2 Contraction of an arbitrary tempered representation

Let me finally consider a general Mackey datum $\delta = (\lambda, \mu)$ and the parabolic subgroup $P_\lambda = M_\lambda A_\lambda N_\lambda$ from section 3.2. Since I have not been able to write down what happens for the contraction of $\mathbf{V}_{M_\lambda}(\mu)$ if it is neither a limit of discrete series (or discrete series) representation nor the one with trivial minimal K_λ-type, I will assume that δ is a nice Mackey datum in the sense of Section 3.3 [10].

Let me consider a carrier Hilbert space $\mathbf{S}^\mu$ for the tempered-irreducible-representation-with-real-infinitesimal-character $\mathbf{V}_{M_\lambda}(\mu)$ of M_λ, σ for the morphism from M_λ to $\mathrm{End}(\mathbf{S}^\mu)$, and let me introduce Hilbert spaces $\mathbf{S}^\mu_t$ for the corresponding representations σ_t of $M_{\lambda,t}$. As I explained in subsection 7.1, we can view all those carrier Hilbert spaces as subspaces of a fixed Fréchet space $\mathbf{E}_\mu$, and we identified a distinguished linear map

$$\mathbf{C}^\mu_t : \mathbf{E}^\mu \to \mathbf{E}^\mu$$

which restricts to a contraction map between $\mathbf{S}^\mu$ and $\mathbf{S}^\mu_t$.

Consider now the vector space $\mathbf{E} = C(K, \mathbf{E}^\mu)$ of continuous functions from K to $\mathbf{E}^\mu$, and endow it with the Fréchet topology of uniform convergence. Pointwise composition with $\mathbf{C}^\mu_t$ defines a linear map

$$\mathbf{C}_t : \mathbf{E} \to \mathbf{E}$$

which sends the subspace $\mathbf{H} \subset \mathbf{E}$ of $\mathbf{S}^\mu$-valued, continuous functions on K which satisfy $f(ku) = \sigma(u)^{-1} f(k)$ for each (k,u) in $K \times (K \cap M_\chi)$, to the subspace $\mathbf{H}_t$ of $\mathbf{S}^\mu_t$-valued, continuous functions on K which satisfy $f(ku) = \sigma_t(u)^{-1} f(k)$ for each (k,u) in $K \times (K \cap M_{t,\chi})$.

Recall from section 4.1 that the representation of G on $\mathbf{H}$ defined by

$$\pi^{comp}_{\lambda,\mu}(g)f = \left[k \mapsto \exp\langle -i\lambda - \rho, \mathbf{a}(g^{-1}k)\rangle \sigma(\mathbf{m}(g^{-1}k)) f\left(\kappa(g^{-1}k)\right)\right]$$

for (g, f) in $G \times \mathbf{H}$, is the compact picture for $\mathbf{M}(\delta)$.

Let me start with the corresponding representation $\pi^{t,comp}_{\lambda,\mu} : G_t \to \mathbf{H}_t$ as before, and define

$$\varpi^{t,comp}_{\lambda,\mu} \quad G \xrightarrow{\varphi_t^{-1}} G_t \xrightarrow{\pi^{t,comp}_{\lambda,\mu}} \mathrm{End}(\mathbf{H}_t).$$

Lemma 7.4. *The linear map $\mathbf{C}_t$ intertwines $\varpi^{t,comp}_{\lambda,\mu}$ and $\pi^{comp}_{\lambda/t,\mu}$.*

Proof. This is but an adaptation of Lemma 4.1. A trivial adaptation of its proof shows that the subgroups $M_{t,\chi}$, $A_{t,\chi}$, $N_{t,\chi}$ used in the definition of $\pi^{t,comp}_{\lambda,\mu}$ are sent by φ_t to the subgroups M_χ, A_χ, N_χ used to define $\pi^{comp}_{\lambda,\mu}$, and that the projections κ_t, $\mathbf{m}_t$, $\mathbf{a}_t$ are related with those for G through

$$\kappa_t(\varphi_t^{-1} g) = \kappa(g);$$

$$\mathbf{a}_t(\varphi_t^{-1} g) = \frac{\mathbf{a}(g)}{t};$$

$$\mathbf{m}_t(\varphi_t^{-1} g) = \mathbf{m}(g).$$

10. Once the extension to all tempered-irreducible-with-real-infinitesimal-character representation is obtained, I hope to be able to just drop this sentence and the results will apply to the full tempered dual.

Because of Lemma 4.2 (or rather the same lemma after a change of notation, and the same proof), we know how the roots of $(\mathfrak{g}_t, \mathbf{a}_\chi)$ evolve with t, and for each $\gamma \in G_t$ we know that

$$\pi^{t,comp}_{\lambda,\sigma}(\gamma) = f \mapsto \left[k \mapsto \exp\langle -i\lambda - \rho, \mathbf{a}_t(\gamma^{-1}k)\rangle \sigma_t(\mathbf{m}_t(\gamma^{-1}k)) f\left(\kappa_t(\gamma^{-1}k)\right)\right].$$

Hence

$$\pi^{t,comp}_{\lambda,\sigma}(\varphi_t^{-1}(g)) = f \mapsto \left[k \mapsto \exp\langle -i\lambda - \rho, \mathbf{a}_t(\left[\varphi_t^{-1}g\right]^{-1}k)\rangle \sigma_t(\mathbf{m}_t(\left[\varphi_t^{-1}g\right]^{-1}k)) f\left(\kappa_t(\left[\varphi_t^{-1}g\right]^{-1}k)\right)\right].$$

And rearranging, we need only recall that $\mathbf{C}_t$ is a contraction map between σ and σ_t to obtain

$$\begin{aligned}
\pi^{t,comp}_{\lambda,\sigma}(\varphi_t^{-1}(g))\left[\mathbf{C}^\sigma_t f\right] &= \left[k \mapsto \exp\langle -i\frac{\lambda}{t} - \rho, t\cdot \mathbf{a}_t(\varphi_t^{-1}\left[g^{-1}k\right])\rangle \sigma_t(\mathbf{m}_t(\left[\varphi_t^{-1}g\right]^{-1}k))\,(\mathbf{C}_t f)\left(\kappa_t(\varphi_t^{-1}\left[g^{-1}k\right])\right)\right]\\
&= \left[k \mapsto \exp\langle -i\frac{\lambda}{t} - \rho, \mathbf{a}(g^{-1}k)\rangle \left\{\mathbf{C}^\sigma_t \sigma(g^{-1}k)\right\} f\left(\kappa(g^{-1}k)\right)\right]\\
&= \mathbf{C}_t\left(\pi^{comp}_{\frac{\lambda}{t},\sigma}(g) f\right),
\end{aligned}$$

so the proof of lemma 7.4 is complete. □

Of course the results of section 7.1 and the description of $\mathbf{M}_0(\delta)$ in section 2.3 mean that

Lemma 7.5. *For each $f \in \mathbf{H}$, there is a limit*[11] *f_0 to $\mathbf{C}_t f$ as t goes to zero. The vector space $\mathbf{H}_0 := \{f_0 \mid f \in \mathbf{H}\}$ carries an irreducible G_0-module with class $\mathbf{M}_0(\delta)$.*

Let me write $\mathbf{S}^\mu_0$ for the subspace of $\mathbf{E}_\mu$ which gathers the limits (in $\mathbf{E}_\mu$) of the $\mathbf{C}^\sigma_t v$, $v \in \mathbf{S}^\mu$, and note that $\mathbf{H}_0 = C(K, \mathbf{S}^\mu_0)$. The next step is to see how the G_0-module structure on $\mathbf{H}_0$ described in section 2.3 emerges from the G-module structure on $\mathbf{H}$ through the contraction process. In section 4.2, I used the track-keeping maps $\alpha_t : G_0 \to G_t$ to show how operators for $\mathbf{M}(\delta)$ converge to operators for $\mathbf{M}_0(\delta)$. Let me proceed in the same way here and set

$$\tilde{\pi}^t = \pi^{t,comp}_{\lambda,\sigma} \circ \alpha_t.$$

Now for each g_0 in G_0, the operator $\tilde{\pi}^t(g_0)$ acts on $\mathbf{H}_t$, and we want to compare it with $\pi_0(g_0)$ which acts on $\mathbf{H}_0$. Point 4. in the statement of Theorem 3.2 above provides a natural way to make the comparison. To see how to prove it, let me come back to discrete series representations for a moment.

In section 5, I used the action of G_t on $\mathfrak{p}$ to build an action on $\mathbf{C}^\infty(\mathfrak{p}, W)$ (beware there is an action on the fibers here). This action is defined on the whole space of continuous functions, and it is by restricting it to the G_t-stable vector subspace of square-integrable solutions of the Dirac equation that we get operators for a discrete series representation. In view of the constructions I recalled in section 6 and 7.1, a finite number of trivial steps extends Lemma 5.6 to the following two facts (in italics) :

There is a family of linear maps $\bar{\sigma}_t : M_{t,\lambda} \to End(\mathbf{E}^\mu)$, weakly continuous w.r.t. the Fréchet topology on $\mathbf{E}^\mu$, such that each $\sigma_t : M_{t,\lambda} \to End(\mathbf{S}^\mu_t)$ is obtained by restricting $\bar{\sigma}_t$

11. The limit is in $\mathbf{E}$ here.

to $\mathbf{E}^\mu$.

In the sequel I will remove the bar and write σ directly for the maps from $\mathbf{E}^\mu$ to itself Because of Lemma 5.6 and its corollary, which extend to limits of discrete series and nonconnected groups with the obvious modifications, these maps will have the following property:

For each g_0, in G_0 and each $f \in \mathbf{E}$, there is a limit (in $\mathbf{E}$) to $\tilde{\sigma}_t(\alpha_t g_0)f$ as t goes to zero. When f belongs to $\mathbf{H}_0$, this limit is $\mu(k)f$.

Now we can use these observations to extend each of the $\pi^{t,comp}_{\lambda,\sigma}$ to all of $\mathbf{E}$, by setting

$$\tilde{\pi}^{t,comp}_{\lambda,\sigma}(\gamma)f = \left[k \mapsto \exp\langle -i\lambda - t\rho, \mathbf{a}_t(\gamma^{-1}k)\rangle \sigma_t(\mathbf{m}_t(\gamma^{-1}k)) f\left(\kappa_t(\gamma^{-1}k)\right)\right],$$

for each γ in G_t and each f in $\mathbf{E}$. Then the linear operator defined on all of $\mathbf{E}$ obtained by setting

$$\pi^t = \tilde{\pi}^{t,comp}_{\lambda,\mu} \circ \alpha_t.$$

extends $\tilde{\pi}^t$.

⋆

Theorem 7.1. *For each g_0 in G_0 and each $f \in \mathbf{E}$, there is a limit (in $\mathbf{E}$) to $\pi_t(g_0)f_t$ as t goes to zero; the limit is $\pi_0(g_0)f_0$.*

Proof. This again extends Theorem 4.1, and the work done since section 5 is enough to have the same strategy work. If $g_0 = (k, v)$, we want to compare

$$\left[\pi^{t,comp}_{\lambda,\sigma}(\exp_{G_t}(v)k)f\right](u) = \left[u \mapsto \exp\langle -i\lambda - t\rho, \mathbf{a}_t((\exp_{G_t}(v)k)^{-1}u)\rangle \sigma_t(\mathbf{m}_t(\exp_{G_t}(v)k)^{-1}u)) f\left(\kappa_t(\exp_{G_t}(v)k)^{-1}u)\right)\right]$$

with

$$\pi_0(g_0)f = \left[u \mapsto e^{i\langle \lambda, Ad(u^{-1})v\rangle} f(k^{-1}u)\right].$$

We first rearrange $\left[\pi^{t,comp}_{\lambda,\sigma}(\exp_{G_t}(v)k)f\right]$ as

$$u \mapsto \exp\langle -i\lambda - t\rho, \mathbf{a}_t\left[k^{-1}u\exp_{G_t}(-Ad(u^{-1})v)\right]\rangle \sigma_t\left(\mathbf{m}_t\left[k^{-1}u\exp_{G_t}(-Ad(u^{-1})v)\right]\right)^{-1} f\left(\kappa_t(k^{-1}u\exp_{G_t}(-Ad(u^{-1})v))\right),$$

and imitate the notation in the proof of Theorem 4.1 by setting $\mathfrak{I}_t = \mathbf{a}_t \circ \exp_{G_t}$, $\mathfrak{K}_t = \kappa_t \circ \exp_{G_t}$.

Then of course $\kappa_t\left(k^{-1}u\exp_{G_t}(-Ad(u^{-1})v)\right) = k^{-1}u\ \mathfrak{K}_t\left[-Ad(u^{-1})v\right]$, and

$$\left[\pi^{t,comp}_{\lambda,\sigma}(k\exp_{G_t} v)f\right] = u \mapsto \exp\langle -i\lambda - t\rho, \mathfrak{I}_t(-Ad(u^{-1})v)\rangle \sigma_t\left(\mathbf{m}_t\left[k^{-1}u\exp_{G_t}(-Ad(u^{-1})v)\right]\right) f\left(k^{-1}u\ \mathfrak{K}_t\left[-Ad(u^{-1})v\right]\right).$$

But

$$\mathbf{m}_t\left[k^{-1}u\exp_{G_t}(-Ad(u^{-1})v\right] = \mathbf{m}_t\left[\exp_{G_t}(-Ad(u^{-1})v)\right].$$

So we can rewrite $\sigma_t(\mathbf{m}_t(\exp_{G_t}(v)k)^{-1}u))$ as

$$\sigma_t \circ \alpha_t \left[1, \mathfrak{M}_t(-Ad(u^{-1})v)\right]$$

with

$$\mathfrak{M}_t = \beta \mapsto \log_{G_t}\left(\mathbf{m}_t(\exp_{G_t}(\beta))\right),$$

a map from $\mathfrak{p}$ to $\mathfrak{m} \cap \mathfrak{p}$. Now, a little playing around with φ_t as in lemma 4.3 shows that

$$\mathfrak{M}_t(\beta) = \frac{1}{t}\log_G\left(\mathbf{m}(\exp_G(t\beta))\right).$$

Just as in lemma 4.3, this means that $\mathfrak{M}_t(\beta)$ goes to the Iwasawa projection of β on $\mathfrak{m} \cap \mathfrak{p}$ along the decomposition $\mathfrak{g} = \mathfrak{k} \oplus (\mathfrak{m} \cap \mathfrak{p}) \oplus \mathfrak{a} \oplus \mathfrak{n}$.

Set $\beta = -Ad(u^{-1})v$. Then we just saw that

$$\left[\pi^{t,comp}_{\lambda,\sigma}(\exp_{G_t}(v)k)f\right] = u \mapsto \underbrace{\exp\left\langle -i\lambda - t\rho, \mathfrak{I}_t(\beta)\right\rangle}\underbrace{(\sigma_t \circ \alpha_t)\left[\mathfrak{M}_t(\beta)\right] f\left(k^{-1}u\ \mathfrak{K}_t\left[\beta\right]\right)}. \tag{7.1}$$

It is time to take the limit, so suppose f is in $\mathbf{E}$ and the above is applied to $f_t = \mathbf{C}_t f$.

• The first underbraced term of course goes to $e^{i\langle -i\lambda, \mathfrak{I}_0(\beta)\rangle} = e^{+i\langle \lambda, Ad(u^{-1})v\rangle}$.

• As for the second underbraced term, because there is a limit to $\mathfrak{M}_t(\beta)$, a straightforward extension of lemma 5.6 shows that there is a distance which defines the topology of $\mathbf{E}^\mu$ and with respect to which all of the $(\sigma_t \circ \alpha_t)\left[\mathfrak{M}_t(\beta)\right]$ (these are operators on $\mathbf{E}^\mu$) are 1-Lipschitz. Rewrite $(\sigma_t \circ \alpha_t)\left[\mathfrak{M}_t(\beta)\right]$ as Σ_t. Then we can rewrite

$$(\sigma_t \circ \alpha_t)\left[\mathfrak{M}_t(\beta)\right] f_t\left(k^{-1}u\ \mathfrak{K}_t\left[\beta\right]\right) - f_0(k^{-1}u) \text{ as:}$$

$$\underbrace{\left[f_t(k^{-1}u) - f_0(k^{-1}u)\right]} + \underbrace{\left[\Sigma_t - Id_E\right]\left[f_t(k^{-1}u) - f_0(k^{-1}u)\right]} + \underbrace{\Sigma_t\left[f_t\left(k^{-1}u\ \mathfrak{K}_t\left[\beta\right]\right) - f_t(k^{-1}u)\right]} + \underbrace{\left[\Sigma_t - Id_E\right]\left[f_0(k^{-1}u)\right]}.$$

The first term goes to zero because it is defined through Lemma 7.5.

The second term goes to zero because of the Lipschitz remark above.

The third goes to zero because $\mathfrak{K}_t(\beta)$ goes to the identity and f_t is defined through pointwise composition of f with a map, namely $\mathbf{C}^\sigma_t$, which is Lipschitz with respect to the given distance on $\mathbf{E}^\mu$ if that distance is chosen appropriately [12].

The last term goes to zero because $\mathfrak{M}_t(\beta)$ is in $\mathfrak{p}$ $f_0(k^{-1}u)$ is $\mathbf{C}^\sigma_t$-invariant and in the corollary to lemma 5.6, it is clear from the proof that when U is in $\mathbf{E}^\mu$, in addition to the convergence (in $\mathbf{E}^\mu$) of $(\sigma_t \circ \alpha_t[k,v])U_t$ to $\mu(k)U_0$ at g_0 fixed, the convergence of $(k,v,x) \mapsto (\sigma_t \circ \alpha_t[k,v])U_t(x)$ to $\mu(k)U_0(x)$ is uniform on compact subsets of $G_0 \times \mathfrak{p}$.

All in all, from (7.1) we see that there is a limit, in $\mathbf{E}$, to $\left[\pi^{t,comp}_{\lambda,\sigma}(\alpha_t[g_0])f_t\right]$ as t goes to zero, and that this limit is $\pi_0(g_0)f_0 = \left(u \mapsto e^{i\langle \lambda, Ad(u^{-1})v\rangle} f_0(k^{-1}u)\right)$, as promised. This is Theorem 7.1. □

12. More precisely, it is enough to choose, in the proof of Lemma 5.6, the compact subsets (A_n) of $\mathfrak{p}$ so that they contain zero and are star domains.

The following statement, a more detailed version of Theorem 3.2, summarizes the contents of sections 4 to 7 :

Theorem 7.2. *Suppose $\delta = (\chi, \mu)$ is a nice Mackey datum.*

Write $\tilde{W}^{\mu}$ for an irreducible K_χ-module of class μ in the case where $\mathbf{V}_{M_\chi}(\mu)$ is a discrete series representation or μ is trivial. In the other case (so that $\mathbf{V}_{M_\chi}(\mu)$ is a limit of discrete series), write $\tilde{W}^{\mu}$ for the finite-dimensional K_χ-module described in section 6.1 as $p_\lambda\left(W_{\lambda+\mu} \otimes F(\mu)\right)$. Set $W^\mu = \tilde{W}^\mu$ in the first case and $W^\mu = p_\lambda\left(W_{\lambda+\mu} \otimes F(\mu)^0\right)$ (see section 6.1) in the second case. Write ν for the cardinal of $M/M^\sharp$ (see section 7.1). Denote by $P_\chi = M_\chi A_\chi N_\chi$ the parabolic subgroup constructed in section 3.

Set $\mathbf{E} = C(K, C(\mathfrak{m}_\chi \cap \mathfrak{p}, \tilde{W}^\mu)^\nu)$, equipped with the Fréchet topology described above. When f is an element of $\mathbf{E}$, define its "value at the origin" as the element of $(W^\mu)^\nu$ obtained by evaluating each component of $f(1_K)$ at the origin of $\mathfrak{m}_\chi \cap \mathfrak{p}$ and projecting the element of $(\tilde{W}^\mu)^\nu$ thus obtained onto $(W^\mu)^\nu$.

Then sections 5 to 7 provide a vector subspace $\mathbf{H} \subset \mathbf{E}$, a map $\pi : G \to End(\mathbf{E})$, and for each $t > 0$ a vector subspace $\mathbf{H}_t \subset \mathbf{E}$ and maps $\pi_t, \Pi_t : G_t \to End(\mathbf{E})$, which have the following properties.

1. Embedding of representations inside the fixed space $\mathbf{E}$. *The vector subspace $\mathbf{H}$ is π-stable, and $(\mathbf{H}, \pi)$ is a tempered irreducible representation of G with class $\mathbf{M}(\delta)$. The vector subspace $\mathbf{H}_t$ is π_t-and Π_t-stable for each $t > 0$, $(\mathbf{H}_t, \pi_t)$ is a tempered irreducible representation of G_t with class $\mathbf{M}_t(\chi, \mu)$, and $(\mathbf{H}_t, \pi_t)$ is a tempered irreducible representation of G_t with class $\mathbf{M}_t(\frac{\chi}{t}, \mu)$.*
2. Existence of a natural family of contraction operators. *For each $t > 0$, there is a unique linear map $\mathbf{C}_t \in End(\mathbf{E})$ which restricts to a contraction map between $(\mathbf{H}, \pi)$ and $(\mathbf{H}_t, \pi_t)$ and preserves the value of functions at the origin. The family $(\mathbf{C}_t)_{t>0}$ is weakly continuous.*
3. Convergence of vectors under the contraction. *For each $f \in \mathbf{E}$, there is a limit (in $\mathbf{E}$) to $\mathbf{C}_t f$ as t goes to zero. Define $\mathbf{H}_0$ as $\left\{ \lim\limits_{t\to 0} \mathbf{C}_t f \;\middle|\; f \in \mathbf{H} \right\}$.*
4. Weak convergence of operators. *Suppose f_0 is in $\mathbf{H}_0$, f is an element of $\mathbf{H}$ with $\lim\limits_{t\to 0} \mathbf{C}_t f = f_0$, and set $f_t = \mathbf{C}_t f$. Then for each g_0 in G_0, there is a limit to $\Pi_t(\alpha_t(g_0)) f_t$ as t goes to zero; the limit depends only on f_0, and belongs to $\mathbf{H}_0$. Call it $\pi_0(g_0)\, f_0$.*
5. The limit produces the appropriate representation of G_0. *The representation $(\mathbf{H}_0, \pi_0)$ of G_0 thus obtained is irreducible unitary, and its equivalence class is $\mathbf{M}_0(\delta)$.*

Remark. The subset of $\widehat{G}$ obtained by considering the classes $\mathbf{M}(\delta)$ with nice δ has full Plancherel measure in $\widehat{G}$; as a result, the above does account for "almost every" tempered representation of G.

When δ is *not* a nice Mackey datum, points 1-4 still hold when $\mathbf{E}$ is chosen appropriately (as a space of functions on K with values in a space of functions on a space smaller than $\mathfrak{m}_\chi \cap \mathfrak{p}$). However, point 5. should disappointingly be replaced the fact that $\mathbf{H}_0$ is a *reducible* unitary representation of G_0; it splits as an (infinite) direct sum of irreducible representations of irreducible G_0-modules in a way that mirrors the decomposition into irreducible K-modules of some $Ind_{K_\chi^\flat}^{K_\chi}(\mu^\flat)$, where $K_\chi^\flat$ and $\mu^\flat$ are using the description of section 6.2. A consequence (using the results of section 2 in Chapter 8) is that $\mathbf{H}_0$ contains with multiplicity one the irreducible representation of G_0 with class $\mathbf{M}_0(\delta)$, and that the corresponding carrier space is that where all the isotypical components for the minimal

K-types of π (and π_0) are to be found.

Remark. Theorem 7.2 is a slightly more precise statement than Theorem 3.2, but does not quite make it obsolete: the variety of possible geometric realizations of a given representation should yield a variety of possible settings to discuss the contraction. Theorem 7.2 is one of them.

8 Concluding remarks

8.1. The contents of sections 4-7 (especially Theorem 7.2) show how, starting from a Hilbert space for $\mathbf{M}(\delta)$, the contraction process wears away everything but a carrier space for $\mathbf{M}_0(\delta)$. However, when Mackey hoped for a result relating the representation theories of G and G_0, his aim was more ambitious and he hoped that new results on G could follow:

> *We feel sure that some such result exists and that a routine if somewhat lengthy investigation will tell us what it is. We also feel that a further study of the apparently rather close relationship between the representation theory of a semisimple Lie group and that of its associated semi-direct product will throw valuable light on the much more difficult semisimple case.*

Higson's constructions certainly throw valuable light on the semisimple case, at least in the case of complex groups, since he shows that the structure of the reduced group $C^\star$-algebras is constant along the deformation from G to G_0, and that an apparently deep fact on the reduced $C^\star$-algebra of G follows from this. I leave it to the reader to decide whether sections 4-7 and the upcoming chapter 8 throw additional light on the semisimple case. But here are some features of the semisimple case which we met on the way:

8.1.1. When realizing a discrete series representation as the space of square-integrable solutions of the Dirac equation for sections of a homogeneous bundle on G/K, lemma 5.5 says that the finite-dimensional subspace carrying the minimal K-type of the representation consists of sections which are entirely determined by their value at the identity coset [13]. I think it is an interesting fact that there are solutions of the Dirac equation which are entirely determined by their value at one given point, that this determines the subspace carrying the minimal K-type, and that these sections are enough to determine the whole representation-theoretic structure of the space of solutions through Vogan's theorem. This is nice, especially because while Theorem 5.1 says these approach constant functions as the contraction is performed, to my knowledge no explicit construction is known for the square-integrable harmonic section with a given value at the identity coset.

8.1.2. This paragraph tries to answer a question put to me by Michèle Vergne. Suppose G/K is hermitian symmetric. Lemma 2.4 shows how the linear map ϕ_t sends G/K, viewed as the G-adjoint orbit of a distinguished elliptic element λ_0, to the G_t-adjoint orbit G_t/K of the same element. Since they are coadjoint orbits, both G/K and G_t/K are symplectic manifolds, and because the Kirillov-Kostant-Souriau form on a coadjoint orbit, or any constant multiple of it, is invariant under the coadjoint action, we see that φ_t provides a symplectic diffeomorphism between G/K and G_t/K if we take care to define the symplectic structures so that they coincide at the tangent space at the identity cosets. Now, we saw on Figure 1 how the G_t-adjoint orbit of λ_0 draws closer and closer to the affine space

13. Of course it is not necessary to bring G_0 into the picture to *prove* lemma 5.5 !

$\lambda_0 + \mathfrak{p}$. This might help discuss a theorem of McDuff [30] which says that G/K and $\mathfrak{p}$ are diffeomorphic as symplectic manifolds, so that G/K admits *global* Darboux coordinates.

The two proofs of this result that I know of [30, 7] use a variation on Moser's homotopy method [31] to obtain a deformation between both symplectic forms. The geometrical setting described in section 2.2 yields a simple explanation for Deltour's proof. Indeed, suppose we start with the symplectic form ω_t on G_t/K defined from the Kirillov-Kostant-Souriau form on $Ad^\star(G_t)\cdot\lambda_0$, and define a symplectic structure Ω_t on $\mathfrak{p}$ as $u_t^\star\omega_t$, multiplied by what is needed in order to have $\Omega_t(0)$ coincide with $\Omega_1(0)$. Then just as it was the case for the riemannian metrics in section 5, lemma 2.2 implies that

$$\Omega_t = \frac{1}{t^2}\left(z_t\right)_\star \Omega_1$$

and of course Ω_t converges to the constant form Ω_0 on $\mathfrak{p}$ as t goes to zero (as before the convergence holds, say, for the topology of uniform convergence of the coefficients in affine coordinates on $\mathfrak{p}$).

The family (Ω_t) is the main ingredient in Deltour's proof of McDuff's theorem: although he does not work with G_t, he uses this very family of symplectic forms and an adaptation of Moser's homotopy method to the noncompact setting to prove that there is an isotopy $(\Psi_t)_{t\in[0,1]}$ of $\mathfrak{p}$, with $\Psi_0 = id_{\mathfrak{p}}$, such that $\Psi_t^\star\Omega_t = \Omega_0$. Very natural indeed if one brings G_t into the picture !

8.2. Here is a list of questions which should get reasonable answers after some work, and to which I hope to come back in the near future:

8.2.1. When an irreducible tempered representation of G with real infinitesimal character is neither in the discrete series nor a limit of discrete series, does it have a realization as a space of functions, or sections of a homogeneous bundle, on G/K ? This would permit to extend Theorem 3.2 to the whole tempered dual.

8.2.2. Suppose $\delta = (\chi, \mu)$ is a Mackey datum and $\mathbf{V}_{M_\chi}(\mu)$ belongs to the discrete series of M_χ. Then J. A. Wolf wrote down in [43] the details for a realization of $\mathbf{M}(\delta)$ as a space of sections of a bundle on $U_\chi := G/(K_\chi A_\chi N_\chi)$ which are square-integrable on each fiber of the natural projection $U_\chi \to K/K_\chi$, and satisfy a partial differential equation gathering the Dirac equations on each of these fibers. It is natural to expect that an easy adaptation of the methods in Section 5 to this realization will lead to another setting for Theorem 3.2 concerning $\mathbf{M}(\delta)$. Does everything go through without any pain ?

8.2.3. The next one was asked by Mackey: how are the (global, distribution) characters of $\mathbf{M}_0(\delta)$ and $\mathbf{M}_\delta$ related ? Since the character of an irreducible representation depend only on its isomorphism class and not on its possible geometric realizations, sections 4 to 7 are not going to be of much help here. A possible direction for understanding this is Kirillov's character formula, which roughly relates the global character to the Euclidean Fourier transform of the Dirac distribution on a coadjoint orbit of G. Rossman proved that the distribution characters of (generic) tempered irreducible representations of G can indeed be exhibited in this way. To my knowledge, this has not been done for G_0 (For a study of the coadjoint orbits of G_0, see [33], however.).

8.2.4. What is the relationship between the Plancherel measure of $\widehat{G}$ and that of $\widehat{G}_0$? This question would call for Harish-Chandra's full work, so it should not be easy. A simpler question would be: how are the Plancherel decompositions of $\mathbf{L}^2(G/K)$ and $\mathbf{L}^2(G_0/K)$ related ? Since this calls only for the spherical principal series, section 4 might be of some help; this would amount to answering question 4.3.3 above. In view of chapter 3 in [16], of section 4 above, and of section 3.6 in [35], this calls for a look at Harish-Chandra's $\mathbf{c}$-function and its behaviour as one "goes to infinity in the Weyl chamber".

8.2.5. When G is a reductive p-adic group, I gather that an anlogue of K and of the tangent space of G/K at the identity, and thus an analogue of G_0, can be defined (but I do not know how). Is there anything to say about the Mackey analogy in that setting ?

Bibliography

[1] M. Atiyah and W. Schmid, *A geometric construction of the discrete series for semisimple Lie groups*, Inventiones mathematicae **54(2)** (1977), 189-192.

[2] P. Baum, A. Connes and N. Higson, *Classifying space for proper actions and K-theory of group $C^\star$-algebras*, Contemporary Mathematics **167**, 241-291.

[3] M. Berger, *A Panoramic View of Riemannian Geometry*, Springer, 2003.

[4] B. Cahen, *On the Dooley-Rice contraction of the principal series*, Note di Matematica **33(2)** (2013).

[5] R. Camporesi, *The Helgason Fourier transform for homogeneous vector bundles over riemannian symmetric spaces*, Pacific Journal of Mathematics **179(2)** (1997), 263-300

[6] A. Connes, *Noncommutative geometry.* Springer Berlin Heidelberg 1994

[7] G. Deltour, *On a generalization of a theorem of McDuff*, Journal of Differential Geometry **93(3)** (2013), 379-400.

[8] A. H. Dooley and J. W. Rice, *On contractions of semisimple Lie groups*, Transactions of the American Mathematical Society **289(1)** (1985), 185-202.

[9] M. Duflo, *Représentations de carré intégrable des groupes semi-simples réels*, Séminaire Bourbaki, vol. 1977/78 Exposés 507-524 (pp. 22-40). Springer Berlin Heidelberg.

[10] M. Duflo, *Opérateurs différentiels bi-invariants sur un groupe de Lie*, Ann. Sci. Ecole Norm. Sup. **10(2)**, (1977), 265-288.

[11] F. Dyson, *Missed Opportunities*, Bulletin of the American Mathematical Society **78(5)** (1972), 635-652.

[12] C. Y. George, *The Mackey analogy for $SL(n,\mathbb{R})$*, Ph.D. Thesis, Pennsylvania State University, 2009.

[13] H. Hecht and W.. Schmid, *A proof of Blattner's conjecture*, Inventiones Mathematicae **31(2)** (1975), 129-154

[14] S. Helgason, *Differential geometry, Lie groups, and symmetric spaces*, Academic press (Vol. 80), New York 1979.

[15] S. Helgason, *Groups & Geometric Analysis: Radon Transforms, Invariant Differential Operators and Spherical Functions.* Academic press, New York 1984.

[16] S. Helgason, *Geometric analysis on symmetric spaces (Vol. 39).* American Mathematical Society, 2008.

[17] N. Higson, *The Mackey analogy and K-theory*, Contemporary Mathematics **449** (2008), 149.

[18] N. Higson, *On the Analogy Between Complex Semisimple Groups and Their Cartan Motion Groups*, Contemporary Mathematics **546** (2011), 137.

[19] E. Inonû and E. P. Wigner, *On the contraction of groups and their representations*, Proc. Nat. Acad. Sci. USA **39** (1953), 510-524.

[20] M. Kashiwara and M. Vergne, *The Campbell-Hausdorff formula and invariant hyperfunctions*, Inventiones mathematicae **47(3)** (1978), 249-272 .

[21] A. A. Kirillov, *Lectures on the Orbit method.* Graduate Studies in Mathematics **64**, Providence, RI: American Mathematical Society

[22] A. W. Knapp, *Representation theory of semisimple groups: An overview based on examples*, 2nd ed., Princeton university press, Princeton 2001.

[23] A. W. Knapp, *Lie groups: beyond an introduction*, 2nd ed., Birkhauser, Boston-Basel-Stuttgart 2001.

[24] A. W. Knapp and G. Zuckerman, *Classification of Irreducible Tempered Representations of Semisimple Groups, Part I.* Annals of Mathematics **116(2)** (1982), 389-455.

[25] A. W. Knapp and G. Zuckerman, *Classification of Irreducible Tempered Representations of Semisimple Groups, Part II.* Annals of Mathematics **116(2)** (1982), 457-492.

[26] J. A. C. Kolk and V. S. Varadarajan, *On the transverse symbol of vectorial distributions and some applications to harmonic analysis*, Indagationes mathematicae **7(1)** (1996), 67-96

[27] G. W. Mackey, *On the analogy between semisimple Lie groups and certain related semi-direct product groups*, in: Lie groups and their representations (Summer School, Bolyai János Math. Soc., Budapest, 1971) Halsted, New York (1975), pp. 339-363.

[28] G. W. Mackey, *The theory of unitary group representations.* Chicago Lectures in Mathematics. University of Chicago Press, Chicago, Ill.-London, 1976.

[29] G. W. Mackey, *Induced representations of locally compact groups.* I. Ann. of Math. (2) 55, (1952). 101-139.

[30] D. McDuff, *The symplectic structures of Kaehler manifolds of nonpositive curvature*, Journal of Differential Geometry **28** (1996), 467-475.

[31] D. McDuff, *Introduction to sympletic topology*, in: Symplectic geometry and topology, ed. Y. Eliashberg and L. Traynor (IAS/Park City mathematical series vol.7), American Mathematical Society (1999), 5-35.

[32] R. Parthasarathy, *Dirac operator and the discrete series.* Annals of mathematics **96** (1972), 1-30.

[33] J. Rawnsley, *Representations of a semidirect product by quantization*, Mathematical Proceedings of the Cambridge Philosophical Society **78(2)** (1975), pp. 345-350.

[34] W. Rossmann, *Kirillov's Character Formula for Reductive Lie Groups*, Inventiones Mathematicae **48** (1978), 207-220.

[35] F. Rouvière *Symmetric spaces and the Kashiwara-Vergne method.* Lecture Notes in Mathematics 2115, Springer Berlin Heidelberg 2014

[36] J. R. Skukalek, *On the Higson-Mackey analogy, Group $C^\star$ algebras, and K-theory*, Ph.D. Thesis, Pennsylvania State University, 2010.

[37] J. R. Skukalek, *The Higson-Mackey analogy for finite extensions of complex semisimple groups*, Journal of Noncommutative Geometry 2015, `http://arxiv.org/abs/1505.06116v1`

[38] P.E. Trapa, *Computing K-type multiplicities in standard modules (after Vogan)*, Notes from an AIM workshop (2005).

[39] D. Vogan, *A Langlands classification for unitary representations*, in: Analysis on homogeneous spaces and representation theory of Lie groups (Okayama-Kyoto 1997). Adv. Stud. Pure Math. **26**, 299-324. Math. Soc. Japan, Tokyo.

[40] D. Vogan, *Branching to a maximal compact subgroup*, in: Harmonic analysis, group representations, automorphic forms and invariant theory, Lect. Notes Ser. Inst. Math. Sci. Natl. Univ. Singap. **12** (2007), pp. 321-401. World Sci. Publ., Hackensack, NJ, 2007.

[41] D. Vogan, *The algebraic structure of the representations of semisimple Lie groups I* Annals of Mathematics **109(1)** (1979), 1-60.

[42] D. Vogan, *Representations of Real Reductive Lie Groups*, Birkhauser, Boston-Basel-Stuttgart 1981.

[43] J. A. Wolf, *Partially harmonic spinors and representations of reductive Lie groups*, Journal of Functional Analysis **15(2)** (1974), pp. 117-154.

Chapter 8

A Mackey-analogy-based proof of the Connes-Kasparov isomorphism for real reductive groups

Contents

Abstract. The Connes-Kasparov conjecture (now a theorem) describes the K-theory of the reduced $C^\star$ algebra of a Lie group in terms of the representation theory of a maximal compact subgroup. When G is a reductive Lie group and G_0 is the Cartan motion group discussed in chapter 7, Alain Connes and Nigel Higson proved in the early 1990s that the Connes-Kasparov conjecture for G is equivalent with the fact that a certain map from the K-theory of the reduced $C^\star$ of G_0 to that of G is an isomorphism, and emphasized the connection with the ideas of George Mackey that have been discussed in Chapter 7.

The conjecture has been verified twice for reductive Lie groups by quite different methods in the late 1980s (Wassermann) and 1990s (Lafforgue), but recently Nigel Higson deepened its connections with the Mackey analogy: he used an elaboration on Mackey's ideas to show that the reduced $C^\star$ algebras of G and G_0 themselves are assembled from identical building blocks, and that the Connes-Kasparov isomorphism is a rather simple reflection of that fact. This chapter shows that his analysis can be extended with the help of the results of Chapter 7; we obtain a Mackey-analogy-based proof of the Connes-Kasparov isomorphism for real (linear connected) reductive groups.

1 Introduction

When G is a second countable locally compact group, there is a natural topology on the unitary dual $\widehat{G}$ (the set of equivalence classes of unitary representations of G). It is known as the Fell topology, and defined by declaring that the closure of a subset $S \subset \widehat{G}$ is the set of equivalence classes of representations whose every matrix element is a limit, for the topology of uniform convergence of compact subsets of G, of matrix coefficients of elements of S.

The Fell topology on $\widehat{G}$ is in general very wild, and studying the topological space $\widehat{G}$ directly is usually difficult. An indirect, often fruitful approach is to study a suitable completion of the convolution algebra $C_c^\infty(G)$ of continuous and compactly supported functions on G. The completion will be a noncommutative algebra, and in the spirit of Connes' noncommutative geometry, the various completions may be thought of as noncommutative replacements for the (not very helpful) space of continuous functions on $\widehat{G}$.

A particularly important completion is the *reduced $C^\star$ algebra* of G. When f is an element of $C_c^\infty(G)$, it can be viewed as a bounded (convolution) operator on the Hilbert space $\mathbf{L}^2(G)$, and that operator has a norm, say $\|f\|$; the reduced $C^\star$ algebra $C_r^\star(G)$ is the completion of $C_c^\infty(G)$ with respect to $\|\cdot\|$. The dual of $C_r^\star(G)$ as a $C^\star$ algebra is not all of $\widehat{G}$, but an important subset $\widehat{G}_r \subset \widehat{G}$ called the reduced dual; it gathers the irreducible representations that are, loosely speaking, necessary to "decompose" the regular representation of G on $\mathbf{L}^2(G)$ into irreducibles.

The *Baum-Connes conjecture* describes the *K-theory* of $C_r^\star(G)$, to be thought of as a "non-commutative" replacement for the Atiyah-Hirzebruch K-theory of $\widehat{G}_r$, in terms of the proper actions of G and a universal example thereof (see [4] for the formulation of the conjecture, and [13] for the K-theory). Because of its generality (it encodes very nontrivial features of both Lie groups and discrete groups, and it has an analogue for groupoids which makes it suitable for the study of foliations), and its deep connections with, and important implications in, geometry, index theory and topology, it has been a leading problem in operator algebra theory for more than thirty years; its study is still a very active field.

The Connes-Kasparov isomorphism. When G is a connected *Lie group*, the conjecture is equivalent to the assertion that the reduced dual of G can be accounted for, at least at the level of K-theory, with the help of Dirac operators. Suppose K is a maximal compact subgroup of G, and $R(K)$ is the representation ring of K − whose underlying abelian group is freely generated by the equivalence classes of irreducible K-modules. Starting from an irreducible K-module, and after going up to a two-fold covering of G and K if necessary, one can build an equivariant spinor bundle on G/K and a natural G-invariant elliptic operator (the Dirac operator) acting on its sections; this operator has an index which can be refined into an element of the K-theory group $\mathcal{K}[C^\star(G)]$. This produces a morphism of abelian groups

$$R(K) \xrightarrow{\mu} \mathcal{K}[C^\star(G)] \tag{1.1}$$

called Dirac induction; the *Connes-Kasparov conjecture* is the statement that μ is an isomorphism.

The case in which G is a *reductive* Lie group is quite special, because then $\widehat{G}_r$ underwent very intense scrutiny between the 1950s and the 1980s, and has been understood completely by Harish-Chandra and a few others who completed his work. As a topological space, $\widehat{G}_r$ is in addition much more reasonable than the reduced duals studied by the full-blown Baum-Connes conjecture: it is (roughly speaking) close to being a real affine variety. Yet the reductive case is a very important one: a major source for the formulation of the Baum-Connes conjecture was the discovery, due to Parthasarathy, Atiyah and Schmid, that Harish-Chandra's deep work on the discrete series of real semisimple Lie groups can be geometrically recovered with the help of Dirac operators on G/K and their $\mathbf{L}^2$-index theory [20, 2]. We shall see at the end of the current paragraph that the reductive case is also the key to the case of general Lie groups.

The Connes-Kasparov conjecture has already been proved; the case of reductive Lie groups has in fact been addressed twice, with two completely different strategies.

Wassermann proved it in a famous short note of 1987 [28]. He used the comprehensive knowledge of $C_r^\star(G)$ extracted by Arthur [3] from Harish-Chandra's monumental work, and his proof consists in an explicit calculation of the right-hand side of (1.1) and the arrow therein. This followed earlier treatment of special cases from the same perspective: the important but simpler case of complex semisimple groups had been covered by Pennington and Plymen [21], and the case of real rank-one groups by Valette [24].

Fifteen years later, Lafforgue obtained a very different proof as a byproduct of a deep study of the actions of groups on Banach spaces and the properties of K-theory ([17]; see also [22]); he found a way to the Baum-Connes isomorphism that is not only very well-suited to reductive Lie groups, but also encompasses reductive p-adic groups as well as some discrete subgroups which had resisted every approach before his. His strategy is almost orthogonal to that of Wassermann, because it succeeds in replacing most of the arsenal of representation theory by a few simple (but far-reaching) facts on the distance to the origin in G/K and on Harish Chandra's elementary spherical function. Lafforgue did not neglect representation theory: he proved that the Connes-Kasparov conjecture actually *implies* that the discrete series can be accounted for with the help of Dirac equations on G/K [18].

Lafforgue's general framework proved flexible enough to be amenable to the extensions needed to prove that (1.1) is also an isomorphism when G is an *arbitrary* (connected [1]) Lie group. Chabert, Etcherhoff and Nest proved that (1.1) is an isomorphism in 2003 [7] by using the fact that an arbitrary Lie group splits as a semidirect product of a reductive and a nilpotent group; the Mackey machine for studying group extensions enabled them to deduce the Connes-Kasparov conjecture from Lafforgue's results on the reductive case.

Lie group deformations and the Mackey analogy. In a relatively recent paper [10], Higson offered a third proof of the Connes-Kasparov conjecture in the special case of complex semisimple Lie groups. It is based on a refomulation of the Connes-Kasparov conjecture in terms of Lie group *deformations* (or *contractions* in the historical terminology of Inonü and Wigner [12]).

Suppose G is a connected Lie group, K is a maximal compact subgroup, and write G_0 for the semidirect product $K \ltimes (\mathfrak{g}/\mathfrak{k})$ associated to the adjoint action of K on $\mathfrak{g}/\mathfrak{k}$ (here $\mathfrak{g}$ and $\mathfrak{k}$ denote the Lie algebras of G and K). Then there is a "continuous" family of groups $\{G_t\}_{t\in\mathbb{R}}$ which interpolates between $G_1 := G$ and G_0 (see section 4 below, as well as section 6 in [10], for precise statements).

This deformation gives rise to a continuous field $\{C^\star(G_t)\}_{t\in\mathbb{R}}$ of algebras; Connes and

1. It is important for their study that the connectivity hypothesis be dropped, but then the Connes-Kasparov conjecture can no longer be formulated with (1.1).

Higson had observed in the 1990s ([6], section 4 in [4] and section 10.β in Chapter 2 of [5]) that the Connes-Kasparov conjecture is equivalent with the fact that this field has *constant* K-theory − in spite of the important difference in structure between $C^\star_r(G_0)$ and $C^\star_r(G)$.

As Connes and Higson insisted, the validity of the Connes-Kasparov conjecture then pointed to an intruiging phenomenon occuring at the level of representation theory: in fact, the reformulation of the Connes-Kasparov conjecture in terms of G_0 echoed enthusiastic observations by G. W. Mackey [19] on a possible relationship between the representation theories of G and G_0 when G is a semisimple Lie group. Mackey had conjectured that there were deep, though surprising, analogies between $\widehat{G}_r$ and $\widehat{G_0}$, and Connes and Higson's observations strongly invited to view the isomorphism (1.1) a simple K-theoretic reflection of these analogies.

Higson took up that idea in more detail in 2008 in the special case of complex semisimple groups; he showed that there is a natural bijection between $\widehat{G}_r$ and $\widehat{G_0}$ in that case, and that analyzing the analogies between $C^\star_r(G_0)$ and $C^\star_r(G)$ in terms of this bijection leads to a proof of the Connes-Kasparov conjecture . This way to the Connes-Kasparov isomorphism takes one through the fine structure of representation theory, but instead of using it for a direct calculation of (1.1) as Wassermann, Pennington and Plymen had successfully done, it expresses the Connes-Kasparov isomorphism as a relatively natural consequence of an easily stated, but actually rather subtle, fact on $\widehat{G}_r$. An appealing feature of Higson's approach from this point of view is that only simple and quite general facts about K-theory are needed, and that no K-theory group need be written down explicitly. His results have been generalized to Lie groups with a finite number of connected components and a complex semisimple identity component in a recent paper of Skukalek [23].

Contents of this note. The representation theory of complex semisimple groups is famous for being much simpler than that of general reductive groups: for instance, the existence of the discrete series (and with it the need for the bulk of Harish-Chandra's work) is a specific feature of the real case. Higson's analysis rests on representation-theoretic facts which, on the surface, may look quite special to the complex case; it is reasonable to wonder whether his strategy can be adapted to real reductive groups.

I have recently exhibited a bijection between $\widehat{G}_r$ and $\widehat{G_0}$ in the real case [1]. In the present note, I use it to prove that Higson's method can indeed be taken up to obtain a proof of the Connes-Kasparov isomorphism for real reductive Lie groups. I should warn the reader that once the results of [1] are brought into the picture, very few ideas need to be added to those in [10] to obtain the results to be described below: I am thus merely going to say how the complex-case-dependent details of Higson's work need to be adapted in order to cover to the case of real groups. As a consequence, I shall use many of the notations and lemmas in [10].

Higson's proof uses Vogan's notion of minimal K-type for the representations of reductive groups [25, 26, 27]. It is assembled from the following four observations.

(a) When G is a complex semisimple Lie group, there is a natural bijection between the reduced duals of G and G_0, and it is compatible with minimal K-types;

(b) To a given set of minimal K-types one can associate a subquotient of the reduced $C^\star$ algebra of G and a subquotient of that of G_0, these subquotients have the same dual (viewed as a topological space), and each is Morita-equivalent with the algebra of continuous functions (vanishing at infinity) on the common dual,

(c) There is a continuous family $(G_t)_{t\in\mathbb{R}}$ of groups interpolating between G and G_0, and given a set of minimal K-types, the road from the corresponding subquotient for $C^\star_r(G_t)$ to the algebra of continuous functions on the common dual varies [2] smoothly with t, even at $t = 0$;

(d) K theory is cohomological in nature, and it is homotopy- and Morita-invariant.

Point (a) is the part most obviously related to the work in [1], and (d) does not depend on the group under consideration; what I am going to argue below is that the part of (a) related to minimal K-types is indeed true, and that in spite of the non-uniqueness of minimal K-types in the real case, the details Higson gave for (b) and (c) can be carried over to real reductive groups without any pain. What is perhaps most surprising in the results of this note is that one should meet no obstacle on the way from complex groups to real groups except those which come from the non-uniqueness of minimal K-types (and which are very easily overcome here).

This note is organized as follows. Section 2 below deals with (a) by proving the assertion about minimal K-types. Section 3 deals with (b); it follows [10] closely − because the subquotients to be defined have a slightly more complicated dual, I merely add some observations furnished by the Knapp-Zuckerman classification of tempered representations and Vogan's results on minimal K-types. Sections 4 and 5 complete the proof of the Connes-Kasparov isomorphism by showing that the results of [10] related with (c) and (d) actually hold without any substantial modification.

2 The Mackey analogy for real groups and minimal K-types

2.1 A bijection between the reduced duals

Let G be a linear connected reductive group. Let me describe the bijection between the tempered dual $\widehat{G}_r$ and $\widehat{G}_0$ used in [1].

Fix a maximal compact subgroup K of G, and let $\mathfrak{p}$ be the orthogonal of $\mathfrak{k}$ in $\mathfrak{g}$ with respect to the Killing form of $\mathfrak{g}$, so that G_0 is isomorphic with the semidirect product $K \ltimes \mathfrak{p}$. For the remainder of this note, fix a maximal abelian subalgebra $\mathfrak{a}$ in $\mathfrak{p}$. When χ is an element of the dual $\mathfrak{a}^\star$, write K_χ for the stabilizer of χ in $\mathfrak{a}^\star$ (for the coadjoint action).

A *Mackey datum* is then a couple (χ, μ) in which χ is in $\mathfrak{a}^\star$ and μ is an irreducible K_χ-module. The set of Mackey data is naturally equipped with an action of the Weyl group $W = W(\mathfrak{g}, \mathfrak{a})$, and thus with an equivalence relation (see section 2.3 in [1]).

If (χ, μ) is a Mackey datum, then one can define a unitary irreducible representation of G_0 by setting

$$\mathbf{M}_0(\chi, \mu) = \mathrm{Ind}_{K_\chi \ltimes \mathfrak{p}}^{G_0} \left[\mu \otimes e^{i\chi}\right]$$

(a realization tailored to our purposes, the "compact picture", will be recalled in section 3.3.2 below).

2. This is admittedly rather vague! I am referring to sections 6.2 and 6.3 of his paper here, and the precise statement to be made below is Proposition B in section 4.

On the other hand, one can define a cuspidal parabolic subgroup $M_\chi A_\chi N_\chi$ of G with the property that the reductive group M_χ admits K_χ as a maximal compact subgroup, and χ can be viewed as an element of $\mathfrak{a}_\chi^\star$ (see [1], section 3.2).

We shall need Vogan's notion of minimal K-type; let me introduce a bit of notation. If G is a reductive Lie group in Harish-Chandra's class and K is a maximal compact subgroup in it, once a maximal torus T is fixed in K and a positive root system is chosen for the pair $(\mathfrak{k}_\mathbb{C}, \mathfrak{t}_\mathbb{C})$, Vogan defines a partial ordering on $\widehat{K}$: if ρ is the half-sum of positive roots with respect to the given ordering, then writing $\tilde{\lambda}$ for the highest weight of an element λ in $\widehat{K}$ and $\langle \cdot, \cdot \rangle$ for the inner product inherited from the Killing form, the K-types can be partially ordered according to the value of $\|\lambda\|_{\widehat{K}} := \langle \tilde{\lambda} + 2\rho, \tilde{\lambda} + 2\rho \rangle$. [3]

Vogan proved that if M is a reductive group and K_M a maximal compact subgroup in M , then by considering the map which associates its minimal K_M-type to an irreducible tempered representation of a reductive group M which has real infinitesimal character, one obtains a bijection onto $\widehat{K_M}$. Returning to a Mackey datum (χ, μ), one can then define an irreducible tempered M_χ-module with real infinitesimal character and minimal K_χ-type μ, say $\mathbf{V}_{M_\chi}(\mu)$. [4] We can then set

$$\mathbf{M}(\chi, \mu) = \mathrm{Ind}_{M_\chi A_\chi N_\chi}^{G} \left[\mathbf{V}_{M_\chi}(\mu) \otimes e^{i\chi + \rho} \right]$$

where ρ is the half-sum of positive roots of $(\mathfrak{g}, \mathfrak{a}_\chi)$ for the ordering used to define N_χ (a realization tailored to our purposes, the "induced picture", will be recalled in section 3.2.4 below).

The starting point for [1] is that $\mathbf{M}_0$ defines a bijection between the set of equivalence classes of Mackey data and the unitary dual of G_0, while $\mathbf{M}$ defines a bijection between the same set and the reduced dual of G.

Remark. In [1], I stayed within the class of linear connected reductive groups with compact center, for fear of using too large a class of reductive groups (since the geometry and Killing-form-induced curvature of G/K played a key role in [1], it was best for that space to be a riemannian symmetric space with negative curvature). However, the definition of the above map between $\widehat{G}_r$ and $\widehat{G_0}$ goes through if the assumption on the center is dropped. The proof that this map provides a bijection between $\widehat{G}_r$ and $\widehat{G}_0$ (which, besides Vogan's results, only uses the Knapp-Zuckerman classification of irreducible tempered representations of linear connected reductive groups with compact center) extends to groups with arbitrary center in an obvious way – one needs only factor both G and G_0 by the vector part of their center. As a consequence, what I just recalled does define a bijection between $\widehat{G}_r$ and $\widehat{G_0}$ when G is linear connected reductive. What I am about to say of the Connes-Kasparov isomorphism will also hold for that class of groups.

2.2 This bijection is compatible with minimal K-types

2.1. Proposition. *When δ is a Mackey datum, the minimal K-types of $\mathbf{M}(\delta)$ and those of $\mathbf{M}_0(\delta)$ coincide.*

3. At some point in the proof of Proposition 2.1 below, I will talk also about the minimal $K \cap M$ types of representations of a reductive subgroup M of G, and for this I shall use the map $\lambda \mapsto \|\lambda\|_{\widehat{K \cap M}}$ associated with the root ordering inherited from those already chosen.

4. Likewise, if M is a reductive group as above, I shall write $\widehat{K_M} \supset \mu \mapsto \mathbf{V}_M(\mu)$ for the inverse of Vogan's map.

Let me first note that as a K-module, $\mathbf{M}_0(\chi, \mu)$ is isomorphic with $\mathrm{Ind}_{K_\chi}^{K}(\mu)$. So what I have to show can be rephrased as follows.

2.1.1. Lemma. *The representations $M(\chi, \mu)$ and $\mathrm{Ind}_{K_\chi}^{K}(\mu)$ have the same minimal K-types.*

Proof. The definition of induced representations shows that $M(\chi, \mu)$ is isomorphic, as a K-module, with $\mathrm{Ind}_{K_\chi}^{K}(\mathbf{V}(\mu)|_{K_\chi})$. Of course that K-module contains $\mathrm{Ind}_{K_\chi}^{K}(\mu)$.

Suppose α is a minimal K-type in $\mathrm{Ind}_{K_\chi}^{K}(\mathbf{V}(\mu)|_{K_\chi})$, but is not a minimal K-type in $\mathrm{Ind}_{K_\chi}^{K}(\mu)$. Then there is μ_1 in $\widehat{K_\chi}$ such that α is a minimal K-type in $\mathrm{Ind}_{K_\chi}^{K}(\mu_1)$, and since α must appear with multiplicity one in $\mathrm{Ind}_{K_\chi}^{K}(\mathbf{V}(\mu)|_{K_\chi})$, μ_1 must appear with multiplicity one in $\mathbf{V}(\mu)|_{K_\chi}$. Because the latter has only one minimal K_χ-type, $\|\mu_1\|_{K_\chi}$ must be greater than $\|\mu\|_K$.

If α were a minimal K-type in $\mathrm{Ind}_{K_\chi}^{K}(\mathbf{V}(\mu_1)|_{K_\chi})$, the representations $\mathrm{Ind}_{K_\chi}^{K}(\mathbf{V}(\mu)|_{K_\chi})$ and $\mathrm{Ind}_{K_\chi}^{K}(\mathbf{V}(\mu_1)|_{K_\chi})$ would have a minimal K-type in common; but lemma 1.2.1.2 below (a reformulation of a theorem by Vogan) says this cannot happen.

Let then α_1 in $\widehat{K}$ be a minimal K-type in $\mathrm{Ind}_{K_\chi}^{K}(\mathbf{V}(\mu_1)|_{K_\chi})$. If α_1 were to appear in $\mathrm{Ind}_{K_\chi}^{K}(\mu_1)$, it would appear in $\mathrm{Ind}_{K_\chi}^{K}(\mathbf{V}(\mu)|_{K_\chi})$, and that cannot be the case because α is already a minimal K-type there.

We end up with α_1 in $\widehat{K}$ and μ_1 in $\widehat{K_\chi}$ such that
– $\|\alpha_1\|_K < \|\alpha\|_K$ and $\|\mu_1\|_{K_\chi} > \|\mu\|_{K_\chi}$
– α_1 is a minimal K-type in $\mathrm{Ind}_{K_\chi}^{K}(\mathbf{V}(\mu_1)|_{K_\chi})$, but it is not a minimal K-type in $\mathrm{Ind}_{K_\chi}^{K}(\mu_1)$.

This seems to trigger an infinite recursion, because the same argument can be used with $(\alpha_1, \mu_1$ instead of (α, μ); however, there are not infinitely many K-types which are strictly lower than α. Thus there cannot be any minimal K-type in $M(\chi, \mu)$ which is not also a minimal K-type in $\mathrm{Ind}_{K_\chi}^{K}(\mu)$: that is Lemma 2.1.1. □

2.1.2. Lemma. *Suppose MAN is a cuspidal parabolic subgroup of G, and μ_1, μ_2 are inequivalent irreducible $K \cap M$-modules. Then the representations $\mathrm{Ind}_{K\cap M}^{K}(\mathbf{V}(\mu_1))$ and $\mathrm{Ind}_{K\cap M}^{K}(\mathbf{V}(\mu_2))$ have no minimal K-type in common.*

Proof. When $\mathbf{V}(\mu_1)$ and $\mathbf{V}(\mu_2)$ are in the discrete series of M, this follows from Theorem 3.6 in [27] (see also Theorem 1 in the announcement [25], and of course [26]). In the other cases, this actually follows from the same result, but I need to give some precisions.

Let me assume that both $\mathbf{V}(\mu_1)$ and $\mathbf{V}(\mu_2)$ are either in the discrete series or nondegenerate limits of discrete series. Then both $\mathrm{Ind}_{MAN}^{G}(\mathbf{V}(\mu_1))$ and $\mathrm{Ind}_{MAN}^{G}(\mathbf{V}(\mu_2))$ are irreducible constituents of some representations induced from discrete series, from a larger parabolic subgroup if need be (see [15], Theorem 8.7). If $\mathrm{Ind}_{M^*A^*N^*}^{G}(\delta_1)$ (with δ_1 in the discrete series of $M^\star$) contains $\mathrm{Ind}_{MAN}^{G}(\mathbf{V}(\mu_1))$ as an irreducible constituent, it contains it with multiplicity one, and the set of minimal K-types are $\mathrm{Ind}_{M^*A^*N^*}^{G}(\delta_1)$ is the disjoint union of the sets of minimal K-types of its irreducible constituents (which are finite in number): see Theorem 15.9 in [14]. If $\mathrm{Ind}_{MAN}^{G}(\mathbf{V}(\mu_1))$ and $\mathrm{Ind}_{MAN}^{G}(\mathbf{V}(\mu_2))$ are constituents of the same representation induced from discrete series, then the lemma follows; if that is not the case we are now in a position to use Vogan's result to the two representations induced from discrete series under consideration (Vogan's disjointness-of-K-types theorem *is* true of reducible induced-from-discrete-series representations).

Now, if $\mathbf{V}(\mu_1)$, in spite of its real infinitesimal character, is neither in the discrete series of M nor a nondegenerate limit of discrete series, there is a smaller parabolic subgroup $M_\star A_\star N_\star$ and a discrete series or nondegenerate limit of discrete series representation ϵ_1 of $M_\star$ such that $\mathrm{Ind}_{K\cap M}^K(\mathbf{V}(\mu_1)) = \mathrm{Ind}_{K\cap M_\star}^K(\epsilon_1)$. If necessary, we can rewrite $\mathrm{Ind}_{K\cap M}^K(\mathbf{V}(\mu_2))$ in an analogous way; then we can use Vogan's result again, after some embeddings in reducible representations induced from discrete series as above if necessary. This proves the lemma. □

The proposition is now established. Let me record here a simple consequence of the very first steps in the proof.

2.2. Corollary. *In each irreducible representation of G_0, every minimal K-type occurs with multiplicity one.*

Proof. Induction is compatible with direct sums, so as a K-module,

$$M(\chi,\mu) = \mathrm{Ind}_{K_\chi}^K\left(\mathbf{V}(\mu)|_{K_\chi}\right) = \mathrm{Ind}_{K_\chi}^K(\mu) \oplus \tilde{M},$$

with $\tilde{M}$ induced from a (quite reducible) K_χ-module; as we saw in Lemma 2.1.1, none of the minimal K-types of $M(\chi,\mu)$ can occur in $\tilde{M}$, and the corollary follows. □

2.3. Remark. C. Y. George considered the tempered dual of $SL(n,\mathbb{R})$ and its Cartan motion group in his thesis [8]. In Chapter 4 there, he describes a set of conjectures which build a bijection between $\widehat{G}_r$ and $\widehat{G_0}_r$ by requiring that it preserve minimal K-types. Proposition 2.1 shows that his conjectures are true and that the bijection they define coincides with that in [1].

3 Some distinguished subquotients of group C*-algebras

3.1 Notations on matrix coefficients

• When λ is an element in the unitary dual of a compact group K, I shall write $d(\lambda)$ for the dimension of the irreducible K-modules with that equivalence class.

• Suppose V_λ is an irreducible K-module with equivalence class λ, and write $\langle\cdot,\cdot\rangle$ for a K-invariant inner product on it. When v is a nonzero vector in V_λ, I set

$$\mathbf{p}_\lambda^v = k \mapsto \langle v, \lambda(k^{-1})v\rangle.$$

Note : When K is connected, the highest-weight vectors cut out a privileged one-dimensional subspace, and any unit highest-weight vector can be chosen for v, yielding a canonical choice for the corresponding matrix element; that is what Higson does in [10]. Because of a slight difference between the situation for real groups and that for complex groups, it will later be useful to choose the vector a bit differently and reassign the name: the rather heavy notation comes from that slight difference.

• I now take up some notations from Higson's paper. Suppose G is a connected unimodular Lie group, K is a compact subgroup, s is a smooth function on K and f is a smooth and compactly supported function on G. Choose a Haar measure on G and define two convolutions between s and f, two smooth and compactly supported functions on G, as

$$s \underset{K}{\star} f = g \mapsto \frac{1}{\mathrm{Vol}(K)} \int_K s(k) f(k^{-1}g)dk \quad , \quad f \underset{K}{\star} s = g \mapsto \int_K f(kg)s(k^{-1})dk.$$

Now suppose K is a compact Lie group, K_1 is a closed subgroup, (V, τ) is an irreducible representation of K with orthonormal basis $\{v_\alpha\}$ and W is a K_1-invariant *irreducible* subspace of V. Write $e_{\alpha\beta}$ for the matrix element $k \mapsto \dim(V)\langle \tau(k)v_\beta, v_\alpha\rangle$ (this is a smooth function on K), and $d_{\alpha\beta}$ for $k \mapsto \dim(W)\langle \tau(k)v_\beta, v_\alpha\rangle$ when v_α and v_β lie in W – the restriction of $d_{\alpha\beta}$ to K_1 is a matrix element of $(W, \tau|_W)$. Then the Schur-Weyl orthogonality relations yield

$$e_{\alpha\beta} \underset{K}{\star} e_{\beta\gamma} = e_{\alpha\gamma}$$

$$d_{\alpha\beta} \underset{K_1}{\star} e_{\beta\gamma} = e_{\alpha\gamma} = d_{\alpha\beta} \underset{K_1}{\star} e_{\beta\gamma} \tag{3.1}$$

(for the second equality, it is to be assumed that v_α, v_β and v_γ lie in W). We also note that $e_{\alpha\beta}(k) = \overline{e_{\beta\alpha}(k^{-1})}$ for all k.

3.2 Subquotients of the reductive group's algebra

3.2.1. The definition.

Let us return to the notations of section 2.1. Define a set

$$\mathcal{K} \subset \left\{ \text{ finite subsets of } \widehat{K} \right\}$$

by declaring that $\mathcal{C}$ is in $\mathcal{K}$ when there is an irreducible tempered representation of G whose set of minimal K-types is $\mathcal{C}$. Note that in this case, $\|\cdot\|_{\widehat{K}}$ takes the same value on all the elements of $\mathcal{C}$.

We are going to associate a subquotient of the reduced C^*-algebra $C^\star_r(G)$ to every element of $\mathcal{K}$. Later on it will be convenient that the family subquotients obtained in this way be associated to an increasing sequence of ideals in $C^\star_r(G)$, so let us choose first a linear ordering

$$\mathcal{K} = \{\mathcal{C}_1, \mathcal{C}_2, ...\}$$

in such a way that
- if the value of $\|\cdot\|_{\widehat{K}}$ on $\mathcal{C}_p$ is (strictly) smaller than that on $\mathcal{C}_q$, then $p < q$,
- if the values of $\|\cdot\|_{\widehat{K}}$ on $\mathcal{C}_p$ and $\mathcal{C}_q$ agree but the number of elements in $\mathcal{C}_p$ is (strictly) larger than that in $\mathcal{C}_q$, then $p < q$.

As soon as we choose choose an arbitrary nonzero vector v_λ in an irreducible K-module with class λ for each λ, and define a matrix element $\mathbf{p}_\lambda = p_\lambda^{v_\lambda}$ as above, we can define a closed ideal in $C^\star_r(G)$ by setting

$$\mathbf{J}[p] = \bigcap_{\lambda \in \mathcal{C}_p} C^\star_r(G)\mathbf{p}_\lambda C^\star_r(G),$$

and a subquotient of $C^\star_r(G)$ by setting

$$\mathbf{C}[p] = (\mathbf{J}[1] + ... + \mathbf{J}[p]) \,/\, (\mathbf{J}[1] + ... + \mathbf{J}[p-1])\,.$$

The dual of $\mathbf{J}[p]$ then gathers the irreducible representations of G whose restriction to K contains every class in $\mathcal{C}_p$. The dual of $\mathbf{C}[p]$ gathers the irreducible representations of G whose set of minimal K-types is exactly $\mathcal{C}_p$.

Remark. If G is a *complex* semisimple Lie group, then every $\mathcal{C}$ in $\mathcal{K}$ contains only one K-type. This leads to some simplifications in Higson's paper and is the reason for several of the notational inconveniences encountered below.

3.2.2 A key lemma. Recall that minimal K-types occur with multiplicity one in irreducible tempered representations of reductive Lie groups. Suppose λ is in $\widehat{K}$ and π is an irreducible tempered representation of G, say on $\mathcal{H}$, containing λ as a minimal K-type. Then for every vector v in the subspace of $\mathcal{H}$ carrying the K-type λ, $\mathbf{p}^v_\lambda$ defines a projection in the multiplier algebra of $\mathbf{C}[p]$, and the projection $\pi(\mathbf{p}^v_\lambda)$ has rank one.

Lemma (Lemma 6.1 in [10]). *Let* $\mathbf{C}$ *be a* C^**-algebra and* $\mathbf{p}$ *be a projection in the multiplier algebra of* $\mathbf{C}$*. If for every irreducible representation* π *of* $\mathbf{C}$ *the operator* $\pi(\mathbf{p})$ *is a rank-one projection, then*

– $\mathbf{CpC} = \mathbf{C}$*;*
– $\mathbf{pCp}$ *is a commutative* C^**-algebra;*
– *the dual* $\widehat{C}$ *is a Hausdorff locally compact space;*
– *The map* $a \mapsto \widehat{a}$ *from* $\mathbf{pCp}$ *to* $\mathcal{C}_0(\widehat{\mathbf{C}})$ *that is defined by*

$$\pi(a) = \widehat{a}(\pi)\pi(\mathbf{p})$$

is an isomorphism of C^**-algebras.*

3.2.3. Some precisions on the subquotient's dual. Part of Lemma 3.2.2 says that the dual of $\mathbf{C}[p]$ is a locally compact Hausdorff space. Let me give some precisions here by recording a simple consequence of Lemma 2.1.2 (Vogan's work) and the Knapp-Zuckerman classification.

I need to introduce some additional notations. Choose an family $P_1, ...P_r$ of nonconjugate cuspidal parabolic subgroups in G, with one element for each conjugacy class of cuspidal parabolic subgroups. Write $M_iA_iN_i$ for the Langlands decomposition of P_i, and K_i for the maximal compact subgroup $K \cap M_i$ in M_i. Anticipating the need for further notation, let me take this opportunity to write $M_i^{\mathfrak{p}}$ for $\exp_G(\mathfrak{m}_i \cap \mathfrak{p})$ and to recall that the Iwasawa map from $K \times (\mathfrak{m}_i \cap \mathfrak{p}) \times \mathfrak{a}_i \times \mathfrak{n}_i$ to G is a diffeomorphism (I use the obvious notation for the Lie algebras here).

Now, define a linear ordering $\widehat{K_i} = \{\lambda_1, \lambda_2, ...\}$ in such a way that if $\|\lambda_n\|_{\widehat{K_i}} < \|\lambda_m\|_{\widehat{K_i}}$, then $n < m$. By *discrete parameter*, I will now mean a couple (i, n) with i in $\{1, ..r\}$ and n in $\mathbb{N}$. Here is the fact on the dual of $\mathbf{C}[p]$ which I shall use:

Lemma. *Let $\mathcal{C}$ be an element of $\mathcal{K}$. There exists a discrete parameter (i_0, n_0) and a subset $\widehat{\mathfrak{a}}[p]$ of $\mathfrak{a}^\star_{i_0}$ which intersects every Weyl group orbit at most once, such that*
- *$\mathbf{V}_{M_{i_0}}(\mu_{n_0})$ is a discrete series or nondegenerate limit of discrete series representation of M_{i_0},*
- *Every irreducible tempered representation of G whose set of minimal K-types is $\mathcal{C}$ is equivalent with exactly one of the*

$$\mathrm{Ind}_{M_{i_0}}^{G}\left[\mathbf{V}_{M_{i_0}}(\mu_{n_0}) \otimes e^{i\chi+\rho}\right], \chi \in \widehat{\mathfrak{a}}[p].$$

- *For every χ in $\widehat{\mathfrak{a}}[p]$, $P_\chi \supset P_{i_0}$ (in fact $M_\chi \supset M_{i_0}$, while $A_\chi \subset A_{i_0}$ and $N_\chi \subset N_{i_0}$).*

This identifies the dual of $\mathbf{C}[p]$ with $\widehat{\mathfrak{a}}[p]$ as a set; we shall see that when $\widehat{\mathfrak{a}}[p]$ is equipped with the topology that it inherits from Euclidean space, the identification becomes a homeomorphism.

Example. It might be useful to recall here that when $G = SL(2, \mathbb{R})$, the spherical principal series representation with nonzero continuous parameter have the same minimal K-types as the (irreducible) spherical principal series representation with continuous parameter zero, but that the nonspherical principal series representation with nonzero continuous parameter have two distinct minimal K-types and the nonspherical principal series representation with continuous parameter zero is reducible. So in that case $\widehat{\mathfrak{a}}[p]$ is either a closed half-line, an open half-line or a single point. In general $\widehat{\mathfrak{a}}[p]$ consists of an open Weyl chamber in a subspace of $\mathfrak{a}$, together with part of one of its walls, and that part is itself stratified analogously until one reaches a minimal dimension (which might be zero, but might not). Note that with the topology inherited from that of the Euclidean space $\mathfrak{a}^\star$, $\widehat{\mathfrak{a}}[p]$ is Hausdorff and locally compact.

3.2.4. An explicit formula for the isomorphism in Lemma 3.2.2.

Suppose a discrete parameter (i_0, n_0), and thus a class $\mathcal{C} = \mathcal{C}_p$, are fixed, and let us start with χ in $\widehat{\mathfrak{a}}[p]$.

In the Hilbert space $\mathcal{V}^{n_0}$ for a representation σ_{n_0} whose equivalence class is $\mathbf{V}_{M_{i_0}}(\mu_{n_0})$, consider the K_{i_0}-isotypical subspace W which corresponds to the K_{i_0}-type μ_{n_0}, and choose a basis $\{v_\alpha\}$ for it.

Let us choose **one** λ_p in $\mathcal{C}_p$. Since p will remain fixed until the end of section 4, I will remove the subscript p from λ. let me use the notation $\mathcal{H}^\chi_{i_0,n_0}$ for a Hilbert space carrying $\mathrm{Ind}_{M_{i_0}}^{G}\left[\mathbf{V}_{M_{i_0}}(\mu_{n_0}) \otimes e^{i\chi+\rho}\right]$ in the usual induced picture: the completion of

$$\left\{\xi : G \underset{\text{comp. supp.}}{\overset{\text{smooth}}{\longrightarrow}} \mathcal{V}^{n_0} \mid \xi(gman) = a^{-i\chi-\rho}\sigma_{n_0}(m^{\mathfrak{p}})^{-1}\xi(g) \;\text{ for }\; (g, m^{\mathfrak{p}}, a, n) \in G \times M^{\mathfrak{p}}_{i_0} \times A_i \times N_i\right\}$$

in the norm associated to the inner product $\langle \xi_1, \xi_2\rangle = \displaystyle\int_K \langle \xi_1(k), \xi_2(k)\rangle_{\mathcal{V}^{n_0}} dk$. I shall write $\pi_{n_0\chi}$ for the usual morphism from G to the unitary group of $\mathcal{H}^\chi_{i_0,n_0}$: $\pi_{n_0\chi}(g)\xi$ is $x \mapsto \xi(g^{-1}x)$ for every (g, ξ) in $G \times \mathcal{H}^\chi_{i_0,n_0}$.

Upon decomposing the λ-isotypical K-invariant component in $\mathcal{H}^{\chi}_{i_0,n_0}$, say V, into K_{i_0}-invariant parts, Frobenius reciprocity says the K_{i_0}-type μ_{n_0} appears exactly once. Fix a K-equivariant identification between the corresponding K_{i_0}-irreducible subspace and W, write $\tilde{v}$ for the vector in the λ-isotypical subspace V which the identification assigns to any v in W. Now, choose an arbitrary v_0 in W (for notational "convenience", I shall assume that it is one of the v_α), and let me introduce a matrix coefficient of our irreducible K-module of type λ as

$$\mathbf{p}_\lambda = \mathbf{p}_\lambda^{v_0} = k \mapsto \langle \tilde{v}_0, \pi_{n_0\chi}(k)\tilde{v}_0\rangle$$

(so although I am suppressing the superscript again, this time I choose a vector in W rather than any nonzero element the λ-isotypical subspace: for complex groups the two choices could be made to coincide because the natural choice was the highest-weight vector which automatically did lie in W, but that is a priori not the case here). Note that the smooth function on K which we just defined actually does not depend on χ.

Henceforth I shall assume that it is this matrix element that is chosen in the definition of $\mathbf{J}[p]$ and $\mathbf{C}[p]$.

The projection $\pi_{n_0\chi}(\mathbf{p}_\lambda)$ has rank one, and to make Lemma 3.2.2 explicit, I shall now *imitate* what Higson did in the complex case, and *use* the calculations he made for the Cartan motion group. A first step is to identify the range of $\pi_{n_0\chi}(\mathbf{p}_\lambda)$. Define a function from K to W as

$$\zeta_{n_0\chi} = k \mapsto \sum_{\alpha=1}^{d(\mu_{n_0})} \langle \pi_{n_0\chi}(k)\tilde{v}_\alpha, \tilde{v}_0\rangle v_\alpha = \frac{1}{d(\lambda)} \sum_{\alpha=1}^{d(\mu_{n_0})} e_{\alpha_0\alpha}(k).$$

We can define a vector in the representation space $\mathcal{H}^{\chi}_{i_0,n_0}$ by extending $\zeta_{n_0\chi}$ to G, setting

$$\xi_{n_0\chi}(km^{\mathfrak{p}}an) = \frac{d(\lambda)^{1/2}}{\mathrm{vol}(K)^{1/2} d(\mu_{n_0})^{1/2}}\, e^{-i\chi-\rho}(a) \sum_{\alpha} \langle \pi_{n_0\chi}(k)\tilde{v}_\alpha, \tilde{v}_0\rangle \cdot \sigma_{n_0}(m^{\mathfrak{p}})^{-1}\,[v_\alpha]\,.$$

Lemma. *The operator $\pi(\mathbf{p}_\lambda)$ on $\mathcal{H}^{\chi}_{i_0,n_0}$ agrees with the orthogonal projection on $\xi_{n_0\chi}$.*

This is easily proved using the formula for the action of K and the inner product on the representation space, and a repeated application of the Schur-Weyl orthogonality relations.

Now put

$$\begin{aligned} \widehat{f}[p] : \qquad & \widehat{\mathfrak{a}}[p] \to \mathbb{C} \\ & \chi \mapsto \int_G f(g) \langle \xi_{n_0\chi}, \pi_{n_0\chi}(g)\xi_{n_0\chi}\rangle \end{aligned} \tag{3.2}$$

as soon as f is a smooth and compactly supported function on G. If $\mathbf{p}_\lambda \underset{K}{\star} f \underset{K}{\star} \mathbf{p}_\lambda = f$, then $\pi_{n_0\chi}(f)$ is proportional to $\pi_{n_0\chi}(\mathbf{p}_\lambda)$, and given the definition of $\pi_{n_0\chi}(f)$, we see that $\pi_{n_0\chi}(f) = \widehat{f}[p](\chi)\ \pi_{n_0\chi}(\mathbf{p}_\lambda)$. This is a first step in making Lemma 3.2.2 explicit, and if I add that $\widehat{f}[p]$ is continuous and vanishes at infinity as a function of χ (this will be obvious from the calculation below), I can summarize this in the following analogue of Lemma

6.10 in [10].

Proposition A. *By associating to any smooth and compactly supported function f on G such that $\mathbf{p}_\lambda \underset{K}{\star} f \underset{K}{\star} \mathbf{p}_\lambda = f$, the element $\widehat{f}[p]$ of $\mathcal{C}_0(\widehat{\mathfrak{a}}[p])$, one obtains a $C^\star$-algebra isomorphism between $\mathbf{p}_\lambda \mathbf{C}[p]\mathbf{p}_\lambda$ and $\mathcal{C}_0(\widehat{\mathfrak{a}}[p])$.*

A consequence which I already mentioned is that the assignation identifies the dual of $\mathbf{C}[p]$ with $\widehat{\mathfrak{a}}[p]$ homeomorphically.

⋆

It will be important later on to have a completely explicit formula for $\widehat{f}[p]$, so I shall now record a closed form for (3.2) which will be useful in section 4.2 below.

Write α_0 for the one α such that $v_0 = v_\alpha$. For a, b in $\{1, ..d(\mu_{n_0})\}$ and χ in $\widehat{\mathfrak{a}}[p]$, set

$$\widehat{f}^{\,p}_{a,b} := \chi \mapsto \frac{\mathrm{Vol}(K)}{d(\lambda)} \int_{N_{i_0}} dn \int_{A_{i_0}} da\, a^{i\chi+\rho} \int_{M^{\mathfrak{p}}_{i_0}} dm \left(f \underset{K_1}{\star} d_{\alpha_0 a} \underset{K_1}{\star} d_{b\alpha_0} \right) (nam) \langle v_b, \sigma_{n_0}(m^{-1}) \, [v_a] \rangle. \tag{3.3}$$

Lemma. *The element $\widehat{f}[p]$ of $\mathcal{C}_0(\widehat{\mathfrak{a}}[p])$ can be expressed as*

$$\widehat{f}[p] = \frac{1}{d(\mu_{n_0})} \sum_{a,b=1}^{d(\mu_{n_0})} \widehat{f}^{\,p}_{a,b}.$$

The proof consists in expanding on (3.2) by simple calculations which closely follow those on page 15 of Higson's paper. We start from the fact that $\langle \xi_{n_0\chi}, \pi_{n_0\chi}(g)\xi_{n_0\chi}\rangle = \int_K \langle \xi_{\mu_{n_0}\chi}(k) \ , \ \xi_{\mu_{n_0}\chi}(g^{-1}k)\rangle dk$; after a change of variables $g \leftarrow g^{-1}k$, and inserting the necessary normalizations to have the $e_{\alpha\beta}$ appear, we find

$$\begin{aligned}
\widehat{f}[p] &= \int_G \left(\int_K f(kg^{-1}) \langle \xi_{\mu_{n_0}\chi}(k), \xi_{\mu_{n_0}\chi}(g) \rangle dk \right) dg \\
&= \int_G \left(\int_K f(kg^{-1}) \frac{1}{\mathrm{Vol}(K)^{1/2} d(\lambda)^{1/2} d(\mu_{n_0})^{1/2}} \sum_\alpha \overline{e_{\alpha_0\alpha}(k)} \langle v_\alpha, \xi_{\mu_{n_0}\chi}(g)\rangle dk \right) dg \\
&= \left(\frac{1}{\mathrm{Vol}(K) d(\lambda)^{1/2} d(\mu_{n_0})} \right)^{1/2} \sum_\alpha \int_G \mathrm{Vol}(K) \, (e_{\alpha\alpha_0} \star f) \, (g^{-1}) \langle v_\alpha, \xi_{\mu_{n_0}\chi}(g)\rangle dg \\
&= \frac{1}{d(\lambda) d(\mu_{n_0})} \sum_{\alpha,\beta} \int_{N_\chi} dn \int_{A_\chi} da \int_{M^{\mathfrak{p}}_\chi} dm \int_K du \, (e_{\alpha\alpha_0} \star f)(n^{-1}a^1 m^{-1} u^{-1}) e_{\alpha_0\beta}(u) \langle v_\alpha, \sigma(m^{-1}) v_\beta\rangle a^{-i\chi-\rho} \\
&= \frac{\mathrm{Vol}(K)}{d(\lambda) d(\mu_{n_0})} \sum_{\alpha,\beta} \int_{N_\chi} dn \int_{A_\chi} da \, a^{-i\chi-\rho} \int_{M^{\mathfrak{p}}_\chi} [e_{\alpha\alpha_0} \star f \star e_{\alpha_0\beta}] \, (n^{-1}a^{-1}m^{-1}) \langle v_\alpha, \sigma(m^{-1}) v_\beta\rangle
\end{aligned}$$

in which the stars are convolutions over K. To relate this to (3.3) we need to have the $d_{\alpha\beta}$ enter the formula in place of the $e_{\alpha\beta}$. We use (3.1) to observe that

$$e_{\alpha\alpha_0} \underset{K}{\star} f \underset{K}{\star} e_{\alpha_0\beta} = d_{\alpha\alpha_0} \underset{K_{i_0}}{\star} \left(e_{\alpha_0\alpha_0} \underset{K}{\star} f \underset{K}{\star} e_{\alpha_0\alpha_0} \right) \underset{K_{i_0}}{\star} d_{\alpha_0\beta} \ ;$$

because of our hypothesis on f this is equal to $d_{\alpha\alpha_0} \underset{K_{i_0}}{\star} f \underset{K_{i_0}}{\star} d_{\alpha_0\beta}$.

To get two convolutions on the right of f instead of one on each side of f, we shorten the above formula by writing $\Gamma^{n_0}_{\alpha\beta}(m)$ for $\langle v_\alpha, \sigma(m^{-1})v_\beta\rangle$, and we use the structure properties of the parabolic subgroup P_{i_0}, along with the fact that χ is K_χ-invariant, thus K_{i_0}-invariant because of the last point in Lemma 3.2.3, to obtain

$$\int_{N_{i_0}} dn \int_{A_{i_0}} da\ a^{-i\chi-\rho} \int_{M^{\mathfrak{p}}_{i_0}} [d_{\alpha\alpha_0} \star f \star d_{\alpha_0\beta}]\,(n^{-1}a^{-1}m^{-1})\Gamma^{n_0}_{\alpha\beta}(m) =$$

$$\int_{N_{i_0}} dn \int_{M^{\mathfrak{p}}_{i_0}} dm\Gamma^{n_0}_{\alpha\beta}(m) \int_{K_{i_0}} dk_1 \int_{K_{i_0}} dk_2 \int_{A_{i_0}} da\ a^{-i\chi-\rho} d_{\alpha\alpha_0}(k_1) f(k_1^{-1}n^{-1}m^{-1}a^{-1}k_2) d_{\alpha_0\beta}(k_2) =$$

$$\int_{N_{i_0}} dn \int_{M^{\mathfrak{p}}_{i_0}} dm\Gamma^{n_0}_{\alpha\beta}(m) \int_{K_{i_0}} dk_1 \int_{K_{i_0}} dk_2 \int_{A_{i_0}} da\ a^{-i\chi-\rho} d_{\alpha\alpha_0}(k_1) f(n^{-1}m^{-1}a^{-1}k_1^{-1}k_2) d_{\alpha_0\beta}(k_2) =$$

$$\int_{N_{i_0}} dn \int_{A_{i_0}} da\ a^{-i\chi-\rho} \int_{M^{\mathfrak{p}}_{i_0}} [f \star d_{\alpha\alpha_0} \star d_{\alpha_0\beta}]\,(n^{-1}a^{-1}m^{-1})\Gamma^{n_0}_{\alpha\beta}(m)$$

(between the first and second line, we used the fact that M_{i_0} centralizes A_{i_0}; between the second and third line, we used the fact that K_{i_0} is contained in M_{i_0}, and thus leaves A_{i_0} invariant and normalizes N_{i_0}, to perform the change of variables $m \leftarrow k_1^{-1}mk_1$, $n \leftarrow k_1^{-1}nk_1$, $a \leftarrow k_1^{-1}ak_1$). The lemma is now proved. $\square$

Remark. Though the calculations are very close to those Higson needed for the Cartan motion group, the presence of a noncompact part in M_{i_0} makes the dependence of the transform $\widehat{f}[p]$ on p less simple than it was for complex groups. But we shall see that the more complicated terms will actually vanish as we perform the contraction to the Cartan motion group.

3.3 Subquotients of the Cartan motion group algebra

3.3.1. I will use the notations of sections 3.1 and 3.2 here. Define a closed ideal in the reduced $C^\star$ algebra $C^\star_r(G_0)$ by setting, as before,

$$\mathbf{J}^0[p] = \bigcap_{\lambda \in \mathcal{C}_p} C^\star_r(G_0)\mathbf{p}_\lambda C^\star_r(G_0),$$

and a subquotient of $C^\star_r(G_0)$ by setting

$$\mathbf{C}^0[p] = (\mathbf{J}[1] + ... + \mathbf{J}[p])\,/\,(\mathbf{J}[1] + ... + \mathbf{J}[p-1])\,.$$

As before, the dual of $\mathbf{C}^0[p]$ gathers the irreducible representations of G_0 whose set of minimal K-types is exactly $\mathcal{C}_p$. Because of Proposition 2.1, we know that the dual of $\mathbf{C}^0[p]$ can be identified *as a set* with $\widehat{\mathfrak{a}}[p]$.

3.3.2. Lemma 3.2.2 is still applicable here of course, and $\mathbf{C}^0[p]$ is thus Morita-equivalent with the algebra of continuous functions, vanishing at infinity, on its dual − viewed as a topological space. To show how nice the correspondence between $\mathbf{C}^0[p]$ and $\mathbf{C}[p]$ is, the next step is to identify the dual of $\mathbf{C}^0[p]$ with $\widehat{\mathfrak{a}}[p]$ not only as a set, but also as a topological space, and to write down an analogue of Proposition 3.2.4. The explicit calculations in Higson's paper are actually sufficient for this, so instead of *imitating* the results of section 5.3 of [10], I shall *invoke* them.

3.3.3. For every χ in $\widehat{\mathfrak{a}}[p]$, Lemma 3.2.3 and Proposition 2.1 show that there is exactly one element $\mu_{p,\chi}$ in $\widehat{K_\chi}$ for which the set of minimal K-types of $\mathbf{M}_0(\chi, \mu_{p,\chi})$ is $\mathcal{C}_p$. Let us write W for the carrier vector space of $\mu_{p,\chi}$ (beware that the meaning of W is no longer the same as in section 3.2)

We shall now use Lemma 5.8 in [10]: we extend χ to $\mathfrak{p}$ by setting it equal to zero on the orthogonal of $\mathfrak{a}_\chi$, we set

$$\widehat{f}(\chi) = \frac{\mathrm{vol}(K)}{d(\lambda)} \int_{\mathfrak{p}} f(x) e^{i\langle \chi, x\rangle} dx$$

as soon as f is a smooth and compactly supported function on G_0, and realize $\mathbf{M}_0(\chi, \mu_{p,\chi})$ as the completion $\mathcal{H}_0$ of

$$\left\{\xi : K \overset{\text{smooth}}{\longrightarrow} V \;\mid\; \xi(gkx) = \mu_{p,\chi}(k)^{-1}\chi(x)^{-1}\xi(g) \text{ for } (k,x,g) \in K_\chi \times \mathfrak{p} \times G_0\right\}$$

in the norm induced by the scalar product between restrictions to K. We write $\pi^0_{\chi,\mu}$ for the G_0-action on induced by left translation.

Now, let us return to semisimple groups for a second. When the λ_p-isotypical subspace V of $\mathcal{H}^\chi_{i_0,n_0}$ is viewed as a K-module, then it must contain the K_χ-type μ exactly once as before because the class of $\pi_{n_0,\chi}$ in $\widehat{G}_r$ is $\mathbf{M}(\chi, \mu_{p,\chi})$; as a consequence, we can obtain a unit vector $\zeta_{p\chi}$ in $\mathcal{H}_0$ by copying the definition of $\zeta_{n_0\chi}$ (the only difference is that what W stands for has changed).

The projection $\pi^0_{\chi,\mu}(\mathbf{p}_\lambda)$ has rank one because of Corollary 2.2. As detailed in Lemma 5.8 of [10], it agrees with the orthogonal projection on $\xi_{p,\chi}$, and the condition $\mathbf{p}_\lambda \underset{K}{\star} f \underset{K}{\star} \mathbf{p}_\lambda = f$ leads to the equality

$$\pi^0_{\chi,\mu}(f) = \widehat{f}(\chi)\; \pi^0_{\chi,\mu}(\mathbf{p}_\lambda).$$

Lemma 3.2.2 then says that is we view $\widehat{\mathfrak{a}}[p]$ as the dual of $\mathbf{C}^0[p]$ and equip it with the corresponding (locally compact and Hausdorff) Fell topology, then $\hat{f}$ becomes a continuous function of χ that vanishes at infinity. But of course $\widehat{f}$ is the ordinary Fourier transform of f, so that it is also a continuous function on $\widehat{\mathfrak{a}}[p]$ when the latter is equipped with the topology inherited from that of Euclidean space. We can summarize the situation with the following statement.

Proposition. *By associating to any smooth and compactly supported function f on G_0 such that $\mathbf{p}_\lambda \underset{K}{\star} f \underset{K}{\star} \mathbf{p}_\lambda$, the element $\widehat{f}$ of $\mathcal{C}_0(\widehat{\mathfrak{a}}[p])$, one obtains a $C^\star$-algebra isomorphism between $\mathbf{p}_\lambda \mathbf{C}^0[p]\mathbf{p}_\lambda$ and $\mathcal{C}_0(\widehat{\mathfrak{a}}[p])$.*

4 Deformation of the reduced C*-algebras and subquotients

This section will be very close to sections 6.2 and 6.3 of [10]; yet, the presence of a noncompact M in the parabolic subgroups which give rise to (3.3) will make it necessary to follow the matrix elements $\langle v_b, \sigma_{n_0}(m^{-1})\,[v_a]\rangle$ through the deformation, and this will lead me to slight changes of notation. Because the setting used in [1] is appropriate for the necessary adaptation, it will be easier for me to define the deformation $\{G_t\}$ to make it coincide with that in [1].

4.1 The continuous field of reduced group $C^\star$ algebras

Let me recall the setting used in [1]: there I worked with a family $\{G_t\}_{t\in\mathbb{R}}$ of groups, together with an isomorphism

$$\varphi_t : G_t \to G$$

for each $t \neq 0$. For every $t \in \mathbb{R}$, the group G_t was *equal* as a topological space with $K \times \mathfrak{p}$, and there is a natural smooth family of measures on $K \times \mathfrak{p}$ giving a Haar measure for each group.

Now let me set

$$\mathcal{G} := \bigsqcup_{t\in\mathbb{R}} G_t.$$

There is a natural bijection

$$\mathcal{G} \to G_0 \cup (G \times \mathbb{R}^\times) \tag{4.1}$$

which is the identity on G_0 and sends $g_t \in G_t$ to $(\varphi_t(g_t), t) \in G \times \mathbb{R}^\times$. Higson recalled on pages 18-19 of [10] how $G_0 \cup (G \times \mathbb{R}^\times)$ has a natural structure of smooth manifold; using the above bijection, we transfer this structure so that $\mathcal{G}$ becomes a smooth manifold, too.

We can consider the reduced $C^\star$ algebra of each G_t once we equip G_t with the Haar measure in the family mentioned above.

Since we chose the manifold structure on $\mathcal{G}$ to make it diffeomorphic with the version used in Higson's paper, lemma 6.13 in [10] shows that the field

$$\{C^\star_r(G_t)\}_{t\in\mathbb{R}}$$

is a continuous field of $C^\star$-algebras. Let me write $\mathcal{C}$ for the $C^\star$ algebra of continuous sections of the restriction of the continuous field $\{C^\star_r(G_t)\}$ to the interval [0,1].

4.2 Subquotients of the continuous field and their spectra

Using the notations of section 3.2, let $\mathcal{J}[p]$ be the ideal $\bigcap_{\lambda\in\mathcal{C}_p} \mathcal{C}\mathbf{p}_\lambda\mathcal{C}$ of $\mathcal{C}$. Define a subquotient of $\mathcal{C}$ by

$$\mathcal{C}[p] = (\mathcal{J}[1] + ...\mathcal{J}[p]) \,/\, (\mathcal{J}[1] + ... + \mathcal{J}[p-1])\,.$$

Higson explained on page 20 of [10] how the dual of $\mathcal{C}[p]$ can be identified with $\widehat{\mathfrak{a}}[p]\times[0,1]$ as a set: the algebra $\mathcal{Z}$ of continuous functions on the closed interval $[0,1]$ lies in the center of the multiplier algebra of $\mathcal{C}[p]$, so that if $\mathcal{Z}^t$ is the subalgebra of functions which vanish at t for t in $[0,1]$, the dual of $\mathcal{C}[p]$ can [5] be identified as a set with the disjoint union over t of the duals of the quotient algebras $\mathcal{C}[p]/(\mathcal{Z}^t\mathcal{C}[p])$. But of course, the algebra $\mathcal{C}[p]/(\mathcal{Z}^t\mathcal{C}[p])$ is isomorphic with the subquotient $\mathbf{C}^t[p]$ of the group algebra $C^\star_r(G_t)$, and we saw that it can be identified with $\widehat{\mathfrak{a}}[p]$.

5. with each t, we need only associate the subset of the dual of $\mathcal{C}[p]$ gathering the representations which restrict to zero on $\mathcal{Z}^t\mathcal{C}[p]$

To proceed further, we need to see how the dual of $\mathcal{C}[p]$ can be identified with $\widehat{\mathfrak{a}}[p] \times [0,1]$ as a topological space. Now, suppose f is a smooth and compactly supported function on $\mathcal{G}$. Write f_t for its restriction to a smooth and compactly supported function on G_t. Write $M_{i_0,t}A_{i_0,t}N_{i_0,t}$ for the cuspidal parabolic subgroup subgroup $\varphi_t^{-1}P$ of G_t, and note that it comes with an ordering on the $\mathfrak{a}_{i_0,t}$-roots and the half-sum ρ_t of positive roots. Suppose $\sigma_{n_0}^t$ is an irreducible representation of $M_{i_0,t}$ with equivalence class $\mathbf{V}_{M_{i_0,t}}(\mu_{n_0})$. Choose a K_{i_0}-equivariant identification between the K_{i_0}-invariant subspace of the Hilbert space for $\sigma_{n_0}^t$ which carries its minimal K_{i_0}-type, and the subspace W of σ_{n_0}. Write (v_a^t) for the basis of that subspace thus associated to (v_a). Gather the transforms considered in sections 3.2 and 3.3, and thus set

$$\widehat{f}^{\,p}_{a,b}(\chi,t) = \begin{cases} \delta_{a,b} \displaystyle\int_{\mathfrak{p}} f_0(x)\chi(x)dx & \text{if } t=0, \\ \displaystyle\int_{N_{i_0,t}} dn_t \int_{A_{i_0}} da_t\; a_t^{i\chi+\rho_t} \int_{M^{\mathfrak{p}}_{i_0},t} dm_t \left(f_t \star d_{\alpha_0 a} \star d_{b\alpha_0}\right)(n_t a_t m_t)\, \langle v_b^t, \sigma_{n_0}^t(m_t^{-1}) \left[v_a^t\right] \rangle & \text{if } t \neq 0. \end{cases} \tag{4.2}$$

for χ in $\widehat{\mathfrak{a}}[p]$. Last, define

$$\widehat{f}^{\,p} = (\chi,t) \mapsto \sum_{a,b=1}^{d(\mu_{n_0})} \widehat{f}^{\,p}_{a,b}(\chi,t), \tag{4.3}$$

a map from $\widehat{\mathfrak{a}}[p] \times [0,1]$ to $\mathbb{C}$.

I will have finished setting the stage for the Connes-Kasparov isomorphism if I check the following fact.

Lemma. *Formula* (4.3) *defines a smooth function on* $\widehat{\mathfrak{a}}[p] \times [0,1]$.

An apparent annoyance to prove this is that the subgroups on which the integrals are taken depend on t. In order to write down (4.2) as an integral over a space which does not depend on t, let me write $\mathfrak{y}_{i_0}$ for $\mathfrak{p} \cap \left(\mathfrak{a}_{i_0} \oplus \mathfrak{m}^{\mathfrak{p}}_{i_0}\right)^{\perp}$ and

$$\beta_t : \mathfrak{y}_{i_0} \to \mathfrak{n}_{i_0,t}$$

for the inverse of $n \mapsto n + n^{\theta}$ (the Cartan involution of $\mathfrak{g}_t$ does not depend on t). Then using exponential coordinates for $\exp_{G_t}(\mathfrak{p})$, remarking that $\mathfrak{m}^{\mathfrak{p}}_{i_0}$ and $\mathfrak{a}_{i_0}$ are $d\varphi_t(0)$-invariant and inserting the fact that ρ_t is none other than $t\rho$ (Lemma 4.2 in [1]), we can rewrite $\widehat{f}^{\,p}_{a,b}(\chi,t)$ as

$$\int_{\mathfrak{y}_{i_0}} dy \int_{\mathfrak{a}_{i_0}} da\; e^{i\langle \chi, a\rangle}\, e^{\langle \rho, ta\rangle} \int_{\mathfrak{m}^{\mathfrak{p}}_{i_0}} dm \left(f \star d_{\alpha_0 a} \star d_{b\alpha_0}\right)\left(\exp_{G_t}(\beta_t y)\exp_{G_t}(a)\exp_{G_t}(m)\right) \langle v_b^t, \sigma_{n_0}^t(\exp_{G_t}(-m)) \left[v_a^t\right]\rangle$$

for $t \neq 0$.

Now I will take up the discussion on page 20 of [10] and use it to remark that f is a smooth, compactly supported function on $\mathcal{G}$ if and only if there is a smooth and compactly supported function F on $K \times \mathfrak{m}^{\mathfrak{p}}_{\chi} \times \mathfrak{a}_{\chi} \times \mathfrak{y}_{\chi} \times \mathbb{R}$ such that

$$F(k,m,a,y) = \begin{cases} f_0(k, m+a+y) \text{ if } t=0; \\ f_t\left(k \exp_{G_t}(m)\exp_{G_t}(a)\exp_{G_t}(\beta_t y)\right) \text{ if } t \neq 0. \end{cases}$$

where f_t is the restriction of f to G_t (compose (4.1) with the proof of Lemma 6.17 in [10]).

Once we insert this, as well as the relationship between $\sigma^t_{n_0}$ and σ detailed in sections 5 and 6 of [1], into (4.2), we can rewrite $\widehat{f}^{\,p}_{a,b}(\chi, t)$ as

$$\int_{\mathfrak{y}_{i_0}} dy \int_{\mathfrak{a}_{i_0}} da e^{i\langle\chi, a\rangle}\ e^{\langle\rho, ta\rangle} \int_{\mathfrak{m}^{\mathfrak{p}}_{i_0}} dm \left[F \star d_{\alpha_0 a} \star d_{b\alpha_0}\right](1, m, a, y, t)\langle v_b, \sigma_{n_0}(\exp_G(-tm))\left[v_a\right]\rangle$$

for $t \neq 0$.

Recall that for fixed p, I considered a $\lambda \in \mathcal{C}_p$ above, and chose a definition for $\mathbf{p}_\lambda$. To check that (4.2) does define a continuous function as soon as f induces an element of $\mathbf{C}[p]$, it is now enough to remember that $\mathfrak{p} = \mathfrak{m}^{\mathfrak{p}}_{i_0} + \mathfrak{a}_{i_0} + \mathfrak{y}_{i_0}$, and to use the Lebesgue (dominated convergence) theorem, the fact that (v_a) is an orthonormal basis, and the Schur-Weyl relations on the matrix elements d. □

Adding this to the results of Section 3, we obtain a statement which sumarizes the rather vague point (c) in the list of the Introduction:

Proposition B. *If f is a smooth and compactly supported function on G_0 and if $\mathbf{p}_\lambda f \mathbf{p}_\lambda = f$, then its transfom $\widehat{f}^{\,p}$ is a continuous function on $\widehat{\mathfrak{a}}[p] \times [0, 1]$, and it vanishes at infinity. This determines an isomorphism of $C^\star$-algebras*

$$\mathbf{p}_\lambda \mathcal{C}[p] \mathbf{p}_\lambda \to \mathcal{C}_0\left(\widehat{\mathfrak{a}}[p] \times [0, 1]\right).$$

5 The Connes-Kasparov isomorphism

We now have gathered all the ingredients needed to prove the Connes-Kasparov conjecture. The argument is not only close to, it is identical with, that in section 7 of [10], and nothing in this section is original in any way.

Evaluation at $t = 0$ and $t = 1$ induce $C^\star$-algebra morphisms from $\mathcal{C}$ to $C^\star_r(G_0)$ and $C^\star_r(G)$, respectively, and in turn these induce two homomorphisms $\alpha_0 : \mathcal{K}(\mathcal{C}) \to \mathcal{K}(C^\star_r(G_0))$ and $\alpha_1 : \mathcal{K}(\mathcal{C}) \to \mathcal{K}(C^\star_r(G_1))$. Because the field $\{C^\star(G_t)\}_{t\in]0,1]}$ (with $t = 0$ excluded) is trivial, α_0 is an isomorphism.

Now, the composition

$$\alpha_1 \circ \alpha_0^{-1} : \mathcal{K}(C^\star(G_0)) \to \mathcal{K}(C^\star(G)) \tag{5.1}$$

is quite relevant to the Connes-Kasparov conjecture, because there is a natural isomorphism between $\mathcal{K}(C^\star(G_0))$ and $R(K)$, and that viewed through this isomorphism, (5.1) is none other than the Connes-Kasparov morphism (1.1). So we shall have reached the aim if we prove that the map α_1, induced by evaluation at $t = 1$, is an isomorphism.

Because $\bigcup\limits_{p\in\mathbb{N}} (\mathcal{J}[1] + ...\mathcal{J}[p])$ is dense in $\mathcal{C}$ and because K-theory commutes with direct limits, we need only prove that evaluation at $t = 1$ yields an isomorphism between $\mathcal{K}(\mathcal{J}[1] + .. + +\mathcal{J}[p])$ and $\mathcal{K}(\mathbf{J}[1] + .. + +\mathbf{J}[p])$ for every p. By standard cohomological arguments, this will be attained if we prove that evaluation at $t = 1$ induces an isomorphism between $\mathcal{K}(\mathcal{C}[p])$ and $\mathcal{K}(\mathbf{C}[p])$.

At this point we recall that for fixed p, the algebras $\mathbf{p}_{\lambda_p}\mathbf{C}[p]\mathbf{p}_{\lambda_p}$ and $\mathbf{C}[p]\mathbf{p}_{\lambda_p}\mathbf{C}[p]$ are Morita-equivalent; therefore the inclusion from $\mathbf{p}_{\lambda_p}\mathbf{C}[p]\mathbf{p}_{\lambda_p}$ to $\mathbf{C}[p]$ induces an isomorphism in K-theory. The problem then reduces to showing that evaluation at $t = 1$ induces an isomorphism between $\mathcal{K}(\mathbf{p}_{\lambda_p}\mathcal{C}[p]\mathbf{p}_{\lambda_p})$ and $\mathcal{K}(\mathbf{p}_{\lambda_p}\mathbf{C}[p]\mathbf{p}_{\lambda_p})$.

We now insert Proposition C and Proposition A to find that between $\mathcal{K}(\widehat{\mathfrak{a}}[p] \times [0,1])$ and $\mathcal{K}(\widehat{\mathfrak{a}}[p])$, evaluation at $t = 1$ *does* induce an isomorphism: this is the homotopy invariance of K-theory. With this the proof that α_1 is an isomorphism is complete.

Remark. It is likely that Skukalek's results which extend Higson's work to almost connected Lie groups with complex semisimple identity component [23] can be used to extend the above proof to almost connected Lie groups with reductive identity component as well (note that these groups may be nonreductive, or may get out of Harish-Chandra's class if they are).

Bibliography

[1] A. Afgoustidis, *How tempered representations of a semisimple Lie group contract to its Cartan motion group* (2015), arXiv:1510.02650.

[2] M. Atiyah and W. Schmid, *A geometric construction of the discrete series for semisimple Lie groups*, Inventiones mathematicae **54(2)** (1977), 189-192.

[3] J. Arthur, *A theorem on the Schwartz space of a reductive Lie group*, Proc Natl Acad Sci U S A. **72(12)** (1975), 4718-9

[4] P. Baum, A. Connes and N. Higson, *Classifying space for proper actions and K-theory of group $C^\star$-algebras*, Contemporary Mathematics **167**, 241-291.

[5] A. Connes, *Noncommutative geometry.* Springer Berlin Heidelberg 1994

[6] A. Connes and N. Higson, *Déformations, morphismes asymptotiques et K-théorie bi-invariante*, C.R.Acad. Sci. Paris, t. **311** (1990), série I, pp. 101-106.

[7] J. Chabert, S. Etcherhoff and R. Nest, *The Connes-Kasparov isomorphism for almost connected groups and for p-adic groups*, **311** (1990), série I, pp. 101-106.

[8] C. Y. George, *The Mackey analogy for $SL(n,\mathbb{R})$*, Ph.D. Thesis, Pennsylvania State University, 2009.

[9] N. Higson, *The Baum-Connes Conjecture*, Proceedings of the Berlin ICM, 1998.

[10] N. Higson, *The Mackey analogy and K-theory*, Contemporary Mathematics **449** (2008), 149.

[11] N. Higson, *On the Analogy Between Complex Semisimple Groups and Their Cartan Motion Groups*, Contemporary Mathematics **546** (2011), 137.

[12] E. Inonû and E. P. Wigner, *On the contraction of groups and their representations*, Proc. Nat. Acad. Sci. USA **39** (1953), 510-524.

[13] G. Kasparov, *Operator K-theory and its applications: Elliptic operator, Group representations, Higher signatures, $C^\star$-extensions*, Proceedings of the Warsaw ICM, 1983.

[14] A. W. Knapp, *Representation theory of semisimple groups: An overview based on examples*, 2nd ed., Princeton university press, Princeton 2001.

[15] A. W. Knapp and G. Zuckerman, *Classification of Irreducible Tempered Representations of Semisimple Groups, Part I.* Annals of Mathematics **116(2)** (1982), 389-455.

[16] A. W. Knapp and G. Zuckerman, *Classification of Irreducible Tempered Representations of Semisimple Groups, Part II.* Annals of Mathematics **116(2)** (1982), 457-492.

[17] V. Lafforgue , *KK théorie bivariante pour les algèbres de Banach et conjecture de Baum-Connes*, Inventiones Mathematicae **149** (2002), 1-95.

[18] V. Lafforgue , *Banach KK theory and the Baum-Connes conjecture*, Proc. ICM Beijing 2002, arXiv:math/0304342.

[19] G. W. Mackey, *On the analogy between semisimple Lie groups and certain related semi-direct product groups*, in: Lie groups and their representations (Summer School, Bolyai János Math. Soc., Budapest, 1971) Halsted, New York (1975), pp. 339-363.

[20] R. Parthasarathy, *Dirac operator and the discrete series.* Annals of mathematics **??** (1972), 1-30.

[21] M. Penington and R. Plymen, *The Dirac operator and the principal series for complex semisimple Lie groups*, J. Functional. Anal. **53** (1983), 269-286.

[22] G. Skandalis, *Progrès récents sur la conjecture de Baum-Connes. Contribution de Vincent Lafforgue.* Séminaire Bourbaki **42** (1999-2000), 105-135

[23] J. R. Skukalek, *The Higson-Mackey analogy for finite extensions of complex semisimple groups*, Journal of Noncommutative Geometry 2015, `http://arxiv.org/abs/1505.06116v1`.

[24] A. Valette, *K-theory for the reduced $C^\star$-algebra of a semisimple Lie group with real rank 1 and finite centre*, Oxford Q. J. Math. **35** (1984), 341-359.

[25] D. Vogan, *Classification of the irreducible representations of semisimple Lie groups*, Proc. Nat. Acad. Sci. USA Vol. **74**, No.7 (july 1977) , pp. 2649-2650.

[26] D. Vogan, *The algebraic structure of the representations of semisimple Lie groups I*, Annals of Mathematics **109(1)** (1979), 1-60.

[27] D. Vogan, *Classifying representations by lowest K-types*, In Applications of group theory in physics and mathematical physics, volume 21 of Lecture Notes in Applied Mathematics, pages 269-288. AMS, Providence, RI, 1985.

[28] A. Wassermann, *Une démonstration de la conjecture de Connes-Kasparov pour les groupes de Lie linéaires connexes réductifs*, C.R. Acad. Sci. Paris, Sér 1 **304** (1987), 559- 562.